AA

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 October 2002 (10.10.2002)

PCT

(10) International Publication Number
WO 02/079218 A1

(51) International Patent Classification[7]: C07H 21/04, C12Q 1/68

(21) International Application Number: PCT/US02/09305

(22) International Filing Date: 27 March 2002 (27.03.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/280,572 30 March 2001 (30.03.2001) US

(71) Applicant: **WYETH** [US/US]; Five Giralda Farms, Madison, NJ 07940 (US).

(72) Inventors: **TREPICCHIO, William, L.**; 21 Abbot Bridge Drive, Andover, MA 01810 (US). **OESTREICHER, Judith, L.**; 125 Main Street, No. 62, Newmarket, NH 03857 (US). **LEONARD, John, P.**; 89 Tennyson Drive, Manchester, NH 03104 (US). **DORNER, Andrew, J.**; 20 Baskin Road, Lexington, MA 02173 (US). **HUNTER, Sharon, E.**; 28 9th Avenue, Haverhill, MA 01830 (US).

(74) Agents: **MANDRA, Raymond, R.** et al.; Fitzpatrick, Cella, Harper & Scinto, 30 Rockefeller Plaza, New York, NY 10112-3801 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/079218 A1

(54) Title: METHODS FOR DIAGNOSING AND TREATING MULTIPLE SCLEROSIS AND COMPOSITIONS THEREOF

(57) Abstract: The present invention is directed to novel methods for diagnosis and prognosis of Multiple Sclerosis by identifying differentially expressed genes. Moreover, the present invention is also directed to methods that can be used to screen test compounds and therapies for the ability to inhibit multiple sclerosis. Additionally, methods and molecule targets (genes and their products) for therapeutic intervention in multiple sclerosis are described.

TITLE

METHODS FOR DIAGNOSING AND TREATING MULTIPLE SCLEROSIS AND COMPOSITIONS THEREOF

[0001] This application claims benefit of U.S. Provisional Patent Application No. 60/280,572, filed March 30,2001.

Field of the Invention

[0002] The present invention is directed to novel methods for diagnosis and prognosis of Multiple Sclerosis by identifying differentially expressed genes. Additionally, methods and molecular targets (genes and their products) for therapeutic intervention in multiple sclerosis are described.

Background of the Invention

[0003] Multiple Sclerosis (MS) is a chronic, often disabling autoimmune disease of the central nervous system that first appears in young adults, with greater occurrence in women than men. MS involves an inflammation of the central nervous system, in which lymphocytes attack myelin and oligodendrocytes to leave behind characteristic lesions or plaques in the brain and spinal cord. The involvement of complex

immunological pathways, in particular systemic T-cell mediated pathways, has presented a challenge to researchers attempting to elucidate the mechanism of the disease.

[0003] Population studies on MS have suggested that MS is a product of genetic susceptibility as compounded by environmental exposure of an unknown nature. Genetic susceptibility of the disease has been borne out by a correlated familial risk, showing the greatest recurrence among monozygotic twins, then among full-siblings and then cousins. Dyment et al. "Genetics of Multiple Sclerosis" Human Molecular Genetics, 6: 1693-1698 (1997). A connection to the environment has further been drawn by several-fold differences in risk between populations in different latitudes, with increasing recurrence in higher latitudes with cooler climates (i.e. northern Europe or Canada). Dyment et al. Population studies however, have been unable to clearly identify all of the genes involved in MS and their complex interaction, nor have these studies been able to provide significant guidance on treatment or therapy for MS.

[0004] The etiology of MS remains unclear, although numerous T-helper cell inflammatory pathways have been implicated in MS pathology. Due to the complex interaction of many genes expressed in these cell-mediated immune pathways, current methodologies have yielded limited success in elucidating the disease. Various methods such as differential display, in situ hybridization, reverse transcriptase-polymerase chain reaction or competitive polymerase chain reaction have been used to study the patterns of gene expression with limited success. Baranzini et al. "Transcriptional

Analysis of Multiple Sclerosis Brain Lesions Reveals a Complex Pattern of Cytokine Expression," *J. Immunol.* 165: 6576-6582 (2000); van Boxel-Dezaire et al. "Decreased Interleukin-10 and Increased Interleukin-12p40 mRNA are Associated with Disease Activity and Characterize Different Disease Stages in Multiple Sclerosis" *Annals of Neurology*, 45: 695-703 (1999)). A study by Whitney and colleagues utilized genetic expression data from one patient to compare MS tissue to non-MS tissue, revealing that about sixty-two genes were differentially expressed, including the Duffy chemokine receptor, interferon regulatory factor-2, and tumor necrosis factor alpha receptor-2. Nonetheless, no method has yet been developed to provide comprehensive data on differentially expressed genes from multiple subjects at multiple stages of development, or through a wide range of geographical regions.

[0005] Furthermore, the nature and variability of MS as expressed in different individuals has proven to be a challenge in characterizing the disease and in providing a prognosis for each patient. Three general stages or types of MS have been characterized: (1) secondary progressive, in which symptoms and disability gradually worsen over time as opposed to having discreet recognizable attacks or relapses; (2) primary progressive, in which there are no attacks but a gradual worsening from the start; and (3) relapsing-remitting, in which there is a clinical worsening of symptoms followed temporary improvement. The present invention therefore addresses these issues by using differentially expressed genes to provide methods for diagnosis and prognosis, and assays for therapeutic intervention.

SUMMARY OF THE INVENTION

[0006] In one embodiment, the invention provides a method of diagnosing a subject with multiple sclerosis by comparing the level of expression of a marker in a sample from a subject, where the marker is selected from the group of markers set forth in Tables 1-5 (for the animal model of multiple sclerosis) or 7-10, to the normal level of expression of the marker in a control sample, where a substantial difference between the level of expression of the marker in the sample from the subject and the normal level is an indication that the subject is afflicted with multiple sclerosis. In a preferred embodiment, the marker corresponds to a transcribed polynucleotide or a portion thereof. Preferably, the marker corresponds to a transcribed polynucleotide or a portion thereof, and the sample is collected from brain tissue or comprises peripheral blood mononuclear cells (PBMCs). In another preferred embodiment, the control sample is from non-involved tissue from the subject. Alternatively, the control sample is from the tissue of a nondiseased subject. In a further preferred embodiment, the level of expression of the marker in the sample differs from the normal level of expression of the marker in a subject not afflicted by a factor of at least two, and in an even more preferred embodiment, the expression levels differ by a factor of at least five.

[0007] In another preferred embodiment, the level of expression of the marker in the sample is assessed by detecting the presence in the sample of a protein corresponding to the marker. In a particularly preferred embodiment, the presence of the protein is

detected using a reagent which specifically binds with the protein. In an even more preferred embodiment, the reagent comprises an antibody or fragments thereof. In another preferred embodiment, the method comprises a marker selected from markers listed in Table 9 or 10. In another preferred embodiment, the level of expression of the marker in the sample is assessed by detecting the presence in the sample of a transcribed polynucleotide or portion thereof, where the transcribed polynucleotide includes the marker. In a particularly preferred embodiment, the transcribed polynucleotide is an mRNA or a cDNA.

[0008] In yet another preferred embodiment, the level of expression of the marker in the sample is assessed by detecting the presence in the sample of a transcribed polynucleotide or a portion thereof which hybridizes with a labeled probe under stringent conditions, wherein the transcribed polynucleotide comprises the marker

[0009] In another preferred embodiment for diagnosing a subject with multiple sclerosis, the level of expression in the sample of each of a panel of markers independently selected from the markers listed in Tables 1-5 or 7-10 is compared with the normal level of expression of the same panel of markers in a control sample, where the level of expression of more than one of the markers is substantially different, relative to the corresponding normal levels of expression of the markers, indicating that the subject is afflicted with multiple sclerosis. In a particularly preferred embodiment, the plurality includes at least five of the markers set forth in Tables 1-5 or 7-10.

[0010] In another embodiment, the invention provides a method of monitoring the progression of multiple sclerosis in a subject, including detecting in a subject sample at a first point in time the expression of marker, where the marker is selected from the group including the markers listed in Tables 1-5 or 7-10, repeating this detection step at a subsequent point in time with the same marker, and detecting a substantial difference between the levels of expression, thus indicating that the subject has progressed to a different stage of multiple sclerosis. In a preferred embodiment, at least 5 markers are selected from the group of markers Tables 1-5 (murine) or 7-10 (human) and combinations thereof. In another preferred embodiment, the marker corresponds to a transcribed polynucleotide or portion thereof, where the polynucleotide includes the marker. In a particularly preferred embodiment, the cells are collected from brain or blood tissue (PBMCs).

[0011] In another embodiment, the invention provides a method of assessing the efficacy of a test compound for inhibiting multiple sclerosis in a subject, including comparing expression of a marker in a first sample obtained from the subject which is exposed to or maintained in the presence of the test compound, where the marker is selected from the group including the markers listed in Tables 1-5 (murine) or 7-10 (human), to expression of the marker in a second sample obtained from the subject, where the second sample is not exposed to the test compound, where a substantially different level of expression of the marker in the first sample relative to that in the second sample is an indication that the test compound is efficacious for inhibiting multiple sclerosis in the subject. In a

preferred embodiment, the first and second samples are portions of a single sample obtained from the subject. In a particularly preferred embodiment, the substantially different level of expression is a lower level of expression in the first sample.

[0012] In another embodiment, the invention provides a method of assessing the efficacy of a therapy for inhibiting multiple sclerosis in a subject, the method including comparing expression of a marker in the first sample obtained from the subject prior to providing at least a portion of the therapy to the subject, where the marker is selected from the group including the markers listed in Tables 1-5 (murine) or 7-10 (human), to expression of the marker in a second sample obtained from the subject following provision of the portion of the therapy, where a substantially different level of expression of the marker in the second sample relative to the first sample, is an indication that the therapy is efficacious for inhibiting multiple sclerosis in the subject. In a preferred embodiment, the substantially different level of expression is a substantially lower level of expression in the second sample. In a particularly preferred embodiment, the method further comprises a step of comparing expression of the marker in a control sample, where a substantially similar level of expression in the second sample, relative to the control sample, is an additional indication that the test compound is efficacious for inhibiting multiple sclerosis.

[0013] In another embodiment, the invention provides a method of screening test compounds for inhibitors of multiple sclerosis in a subject, the method including obtaining a sample including cells from a subject,

separately maintaining aliquots of the sample in the presence of a plurality of test compounds, comparing expression of a marker in each of the aliquots, where the marker is selected from the group including the markers listed in Tables 1-5 (murine) or 7-10 (human), and selecting one of the test compounds which induces a substantially different level of expression of the marker in the aliquot containing that test compound, relative to other test compounds. In a particularly preferred embodiment, the substantially different level of expression is a substantially lower level of expression. In an alternative preferred embodiment, the substantially different level of expression is a substantially enhanced level of expression.

[0014] In another embodiment, the invention provides a kit for diagnosing a subject with multiple sclerosis, including reagents for assessing expression of a marker selected from the group including the markers listed in Tables 1-5 (murine) or 7-10 (human).

[0015] In another embodiment, the invention provides a kit for diagnosing multiple sclerosis in a subject, the kit including a nucleic acid probe where the probe specifically binds with a transcribed polynucleotide corresponding to a marker selected from the group including the markers listed in Tables 1-5 (murine) or 7-10 (human).

[0016] In another embodiment, the invention provides a kit for assessing the suitability of each of a plurality of compounds for inhibiting multiple sclerosis, the kit including a plurality of compounds and a reagent for assessing expression of a marker

selected from the group including the markers listed in Tables 1-5 (murine) or 7-10 (human).

[0017] In another embodiment, the invention provides a kit for diagnosing a subject with multiple sclerosis, the kit including an antibody which specifically binds with a protein corresponding to a marker selected from the group including the markers listed in Tables 1-5 (murine) or 7-10 (human).

[0018] In another embodiment, the invention provides a method of modulating the level of expression of a marker selected from the markers listed in Tables 1-5 (murine) or 7-10 (human), the method comprising providing to diseased cells of the subject an antisense oligonucleotide complementary to a polynucleotide corresponding to the marker.

[0019] In yet another embodiment, the invention provides a method of modulating the level of expression of a marker selected from the markers listed in Tables 1-5 (murine) or 7-10 (human), the method comprising providing to diseased cells of a subject a protein. In a particularly preferred embodiment, the invention further provides a vector which comprises a polynucleotide encoding the protein.

[0020] In another embodiment, the invention provides a method of modulating a level of expression of a marker selected from the markers listed in Tables 1-5 (murine) or 7-10 (human), where the method comprises providing to diseased cells of a subject an antibody. In a particularly preferred embodiment, the method further comprises a therapeutic moiety conjugated to the antibody.

[0021] In another preferred embodiment, the invention provides a method of localizing a therapeutic moiety to diseased tissue of a subject comprising exposing the tissue to an antibody which is specific to a protein encoded by a marker listed in Tables 1-5 (murine) or 7-10 (human).

[0022] In another preferred embodiment, the present invention provides a method of screening for a test compound capable of modulating the activity of a protein encoded from a marker listed in Tables 1-5 (murine) or 7-10 (human), said method comprising combining said protein and test compound, and determining the effect of said test compound on the therapeutic efficacy of said protein.

[0023] In yet another preferred embodiment, the present invention provides a method of screening for a bioactive agent capable of interfering with the binding of a protein or a fragment thereof and an antibody which binds to said protein or fragment thereof, where the method combines a protein or fragment thereof, a bioactive agent and an antibody which binds to the protein or fragment thereof, wherein the method further includes determining the binding of the protein or fragment thereof and the antibody.

[0024] In another preferred embodiment, the present invention provides an antibody which specifically binds to a protein encoded from a marker listed in Tables 1-5 (murine) or 7-10 (human). In particularly preferred embodiment, the antibody is monoclonal and humanized.

[0026] In yet another preferred embodiment, the present invention provides a peptide encoded from markers listed in Tables 1-5 (murine) or 7-10 (human). Furthermore, the present invention is also directed to a composition comprising the peptide.

[0027] In an alternative embodiment, the present invention provides a composition capable of modulating an immune response in a subject, where the composition comprises a protein encoded from a marker listed in Tables 1-5 (murine) or 7-10 (human) and a pharmaceutically acceptable carrier.

[0028] In yet another embodiment, the present invention provides a biochip comprising a panel of markers selected from the group of markers listed in Tables 1-5 (murine) or 7-10 (human). Furthermore, in a particularly preferred embodiment, the markers for a biochip may be selected for subjects suspected of having multiple sclerosis from different stages of the disease: secondary progressive, primary progressive, relapsing-remitting. In a still another embodiment, the markers may be selected for a subject which is from a higher-risk geographical region.

[0029] Other features and advantages of the invention will be apparent from the following detailed description and claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0030] FIGURE 1A through 1D depict the relative expression levels of four cytokines, IL-10(FIGURE 1A; $p<0.4$), IL-8(FIGURE 1B; $p<0.027$), IL-12p35(FIGURE 1C;

$p<0.028$) and IL-1β(FIGURE 1D; $p<0.011$), in MS-afflicted subjects versus nondiseased subjects, as measured in Example 2(D) below. The sample size was 10 MS-afflicted subjects and 10 nondiseased subjects.

[0031] FIGURE 2 depicts the results of the class predictor used in Example 2(E), wherein a diagnosis of "normal" is indicated as positive and "MS-afflicted" are indicated as negative. The model was based on markers listed in Table 10, as expressed in human blood samples.

DETAILED DESCRIPTION OF THE INVENTION

[0032] The present invention provides methods for diagnosis and prognosis evaluation for multiple sclerosis (MS) in subjects, as well as methods and molecular targets for therapeutic intervention.

[0033] In one aspect of the invention, the expression levels of genes are determined in a particular patient sample for which either diagnosis or prognosis information is desired. The level of expression of a number of genes simultaneously provides an expression profile, which is essentially a "fingerprint" of the activity of a gene or plurality of genes that is unique to the state of the cell. Comparison of relative levels of expression have been found to be indicative of the presence of multiple sclerosis, and as such permits for diagnostic and prognostic analysis. By comparing relative expression profiles of multiple sclerosis tissue in known different states (i.e. secondary progressive vs. primary progressive vs. relapsing-remitting), information regarding which genes are important (including both up- and down-regulation of genes) in each of these states is obtained. The

identification of gene markers that are differentially expressed in diseased versus non-diseased tissue, as well as differential expression resulting in different prognostic outcomes, allows the use of this invention in a number of ways. For example, the evaluation of a particular treatment regime may be evaluated: will a particular drug act to improve the long-term prognosis in a particular patient? The discovery of these differential expression patterns for individual genes allows for screening of drug candidates with an eye to mimicking or altering a particular expression pattern; for example, screening can be done for drugs that will alter the MS differential expression pattern or convert a poor prognosis pattern to a better prognosis pattern. This may be done by making biochips comprising sets of the significant MS genes, which can then be used in these screens. These methods can also be done on the protein basis; that is protein expression levels of the MS-associated proteins can be evaluated for diagnostic and prognostic purposes or to screen test compounds. In addition, the markers can be administered for gene therapy purposes, including the administration of antisense nucleic acids, or proteins (including antibodies and other modulators thereof) administered as therapeutic drugs.

[0034] Moreover, in a preferred embodiment, the relative expression levels are measured from human samples; however, as will be appreciated by those in the art, expression levels of markers from other organisms may be useful in animal models of disease and drug evaluation; thus other markers are provided, from vertebrates, including mammals, including rodents (mice, hamsters, rats, guinea pigs, etc), primates, farm animals (including sheep, goats, pigs, cows,

horses, etc). Markers from other organisms may be obtained using the techniques outlined below.

[0035] The present invention is based, at least in part, on the identification of a number of genetic markers, set forth in Tables 1-5 (mouse) and 7-10 (human), which are differentially expressed between diseased samples (MS-associated) and non-diseased samples. Using a murine model of experimental autoimmune encephalitis ("EAE") as an analogy to human MS, a panel of 11,000 known murine genes was screened for expression in diseased versus non-diseased tissue from twelve different mice afflicted with the disease (see Example 1). Those genes with statistically substantial differences between the diseased and normal tissues are identified in Tables 1-2. This differential expression was observed either as an increase in expression (Table 1), or a decrease in expression (Table 2). In addition, to narrow the subset of diseased-related, immune-mediated genes, diseased cells were stimulated *in vitro* with a growth protein and with an inhibitory compound, to yield 6 genes which were differentially expressed: CAPN12 (calpain 12), MT1 (metallothioneln 1), MYO1F (myosin 11), TLN (talin), UNK_AA117532 (EST), and UNK_AA645990 (EST).

[0036] Using the murine model as a springboard, a panel of 12,000 human genes was likewise screened for differential expression between diseased samples of the brain and blood versus non-diseased samples (either from noninvolved brain tissue or from non-diseased subjects) (see Example 2). Genes that were differentially up- or down- regulated in MS blood (peripheral blood mononuclear cells, or "PBMCs")

samples are shown in Table 7, while genes that were differentially up- or down-regulated in MS brain samples are shown in Table 8. Furthermore, comparison of the differentially expressed genes between blood and brain samples yielded 181 genes in common. Further comparison of the 181 genes to the murine genes that were differentially regulated *in vitro* yielded six genes (shown in Table 9 with their accession numbers): EEF1D, PIM2, PRDX2, SEC24C, UNK_AJ24 AND XIP.

[0037] Included among the genes used to screen diseased versus non-diseased tissue in the murine panel were two genes known in the art to be implicated in EAE: TNF-α and IFN-β. These genes served as an internal control. Each of these genes were found to be substantially increased in expression in EAE cells as opposed to non-diseased cells, thus validating the method as a means for identifying significant genes involved in EAE pathology. Correspondingly, the genes which are known in the art to be linked to MS (listed in Table 6) may also serve as validation in expression studies for MS. Moreover, the differentially regulated genes of the invention, as listed in Tables 1-5 (murine) or 7-10 (human), have not been previously associated with EAE or multiple sclerosis.

[0038] Accordingly, the present invention pertains to the use of the genes set forth in Tables 1-5 (murine) or 7-10 (human), the corresponding mRNA transcripts, and the encoded polypeptides as markers for the presence or risk of development of MS. These markers are further useful to correlate the extent and/or severity of disease. In particular, the present invention is directed to the genes set forth in Table 9 (genes which were shown to be differentially regulated

in both murine and human models of MS), and in Table 10, the genes which were used in the MS class predictor of Example 2(E). Panels of the markers can be conveniently arrayed on solid supports, i.e. biochips for use in kits. Markers can also be useful for assessing the efficacy of a treatment or therapy of MS.

[0039] In one aspect, the invention provides markers whose level of expression, which signifies their quantity or activity, is correlated with the presence of MS. The markers of the invention may be nucleic acid molecules (e.g., DNA, cDNA or mRNA) or peptide(s). Preferably the invention is performed by detecting the presence of a transcribed polynucleotide or a portion thereof, wherein the transcribed polynucleotide comprises the marker. Alternatively, detection may be performed be detecting the presence of a protein which corresponds to the marker. The markers of the invention are either increased or decreased in quantity or activity in MS tissue as compared to non-diseased tissue. For example, the gene designated 'SAA3' is increased in expression level in murine EAE cells, relative to control cells, while the gene designated 'E_TC36651_s" is decreased in expression level in murine EAE cells, relative to control cells. Both the presence of increased or decreased mRNA for these genes (and for other genes set forth in Tables 1-5 and 7-10), and also increased or decreased levels of the protein products of these genes (and other genes set forth in Tables 1-5 and 7-10) serve as markers for either EAE or MS. Preferably, increased or decreased levels of the markers of the invention are increases and decreases of a magnitude that are statistically substantial as compared to appropriate control samples (i.e., non-

involved tissue or from non-diseased subjects.) In particularly preferred embodiments, the marker is increased or decreased relative to control samples by at least 2-, 3-, 4-, 5-, 6-, 7-, 8-, 9-, or 10- fold or more. Similarly one skilled in the art will be cognizant of the fact that a preferred detection methodology is one in which the resulting detection values are above the minimum detection limit of the methodology.

[0040] Detection and measurement of the relative amount of a nucleic acid or peptide marker of the invention may be by any method known in the art (see, i.e., Sambrook, J., Fritsh, E.F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989), and *Current Protocols in Molecular Biology*, eds. Ausubel et al, John Wiley & Sons (1992)). Typical methodologies for detection of a transcribed polynucleotide include RNA extraction from a cell or tissue sample, followed by hybridization of a labeled probe (i.e., a complementary nucleic acid molecule) specific for the target RNA to the extracted RNA and detection of the probe (i.e. Northern blotting). Typical methodologies for peptide detection include protein extraction from a cell or tissue sample, followed by hybridization of a labeled probe (i.e., an antibody) specific for the target protein to the protein sample, and detection of the probe. The label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Detection of specific peptide(s) and nucleic acid molecules may also be assessed by gel electrophoresis, column chromatography, direct sequencing, or quantitative PCR (in the case of nucleic

acid molecules) among many other techniques well known to those skilled in the art.

[0041] In certain embodiments, the genes themselves (i.e., the DNA or cDNA) may serve as markers for MS. For example, the absence of nucleic acids corresponding to a gene (i.e. a gene from Table 2) such as by deletion of all or part of the gene, may be correlated with disease. Similarly an increase of nucleic acid corresponding to a gene (i.e. a gene from Tables 1-5 and 7-10), such as by duplication of the gene, may also be correlated with disease.

[0042] Detection of the presence or number of copies of all or a part of a marker gene of the invention may be performed using any method known in the art. Typically, it is convenient to assess the presence and/or quantity of a DNA or cDNA by Southern analysis, in which total DNA from a cell or tissue sample is extracted, is hybridized with a labeled probe (i.e. a complementary DNA molecules), and the probe is detected. The label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Other useful methods of DNA detection and/or quantification include direct sequencing, gel electrophoresis, column chromatography, and quantitative PCR, as is known by one skilled in the art.

[0043] The invention also encompasses nucleic acid and peptide molecules which are structurally different from the molecules described above (i.e. which have a slight altered nucleic acid or amino acid sequence), but which have the same properties as the molecules above (e.g., encoded amino acid sequences, or which are changed only

in nonessential amino acid residues). Such molecules include allelic variants, and are described in greater detail in subsection I.

[0044] In another aspect, the invention provides markers whose quantity or activity is correlated with different stages of MS: secondary progressive, primary progressive, relapsing-remitting. These markers are either increased or decreased in quantity or activity in MS tissue in a fashion that is either positively or negatively correlated with the degree of severity of the MS. A method of monitoring progression of MS in subjects may be devised by detecting a substantial difference between the levels of expression in a diseased subject at different points in time. The subsequent level of expression may further be compared to different expression profiles of various MS stages to confirm whether the subject has a matching profile. In yet another aspect, the invention provides markers whose quantity or activity is correlated with a risk in a subject for developing MS. For instance, the markers may be selected for higher-risk geographical regions. These markers are either increased or decreased in activity or quantity in direct correlation to the likelihood of the development of MS in a subject.

[0045] Each marker may be considered individually, although it is within the scope of the invention to provide combinations of two or more markers for use in the methods and compositions of the invention to increase the confidence of the analysis. In another aspect, the invention provides panels of the markers of the invention. In a preferred embodiment, these panels of markers are selected such that the markers within any one panel share certain features (see Example

2(E)). For example, the markers of a first panel may each exhibit a decrease in quantity or activity in MS tissue as compared to samples from non-involved samples from the same subject or tissue from a non-diseased subject. Similarly, different panels of markers may be composed of markers from different tissues (i.e., blood (Table 7) or brain tissue (Table 8)), or may represent different components of an MS stage or type (i.e., secondary progressive, primary progressive, relapsing-remitting in MS humans; or onset, peak and recovery in EAE mice). Panels of the markers of the invention may be made by independently selecting markers from any of Tables 1-5 for mice, and Tables 7-10 for humans, and may further be provided on biochips, as discussed below.

[0046] It will be appreciated by one skilled in the art that the panels of markers of the invention may conveniently be provided on solid supports, as a biochip. For example, polynucleotides may be coupled to an array (e.g., a biochip using GeneChip® for hybridization analysis), to a resin (e.g., a resin which can be packed into a column for column chromatography), or a matrix (e.g. a nitrocellulose matrix for northern blot analysis). The immobilization of molecules complementary to the marker(s), either covalently or noncovalently, permits a discrete analysis of the presence or activity of each marker in a sample. In an array, for example, polynucleotides complementary to each member of a panel of markers may individually be attached to different, known locations on the array. The array may be hybridized with, for example, polynucleotides extracted from a brain sample from a subject. The hybridization of polynucleotides from the sample with the array at any location on the

array can be detected, and thus the presence or quantity of the marker in the sample can be ascertained. In a preferred embodiment, an array based on a biochip is employed. Similarly, Western analyses may be performed on immobilized antibodies specific for different polypeptide markers hybridized to a protein sample from a subject.

[0047] It will also be apparent to one skilled in the art that the entire marker protein or nucleic acid molecule need not be conjugated to the biochip support; a portion of the marker or sufficient length for detection purposes (i.e., for hybridization), for example a portion of the marker which is 7, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 100 or more nucleotides or amino acids in length may be sufficient for detection purposes.

[0048] The nucleic acid and peptide markers of the invention may be isolated from any tissue or cell of a subject. In a preferred embodiment, the tissue is brain tissue. However, it will be apparent to one skilled in the art that other tissue samples, including bodily fluids such as blood, may also serve as sources from which the markers of the invention may be assessed. The tissue samples containing one or more of the markers themselves may be useful in the methods of the invention, and one skilled in the art will be cognizant of the methods by which such samples may be conveniently obtained, stored and/or preserved.

[0049] Several markers were known prior to the invention to be associated with MS and are provided in Table 6. These markers are not to be considered as

markers of the invention. However, these markers may be conveniently be used in combination with the markers of the invention (Tables 1-5 and 7-10) in the methods, panels and kits of the invention.

[0050] In another aspect, the invention provides methods of making an isolated hybridoma which produces an antibody useful for diagnosing a patient with MS. In this method, a protein corresponding to a marker of the invention is isolated (e.g., by purification from a cell in which it is expressed or by transcription and translation of a nucleic acid encoding the protein *in vivo* or *in vitro* using known methods). A vertebrate, preferably a mammal such as a mouse, rabbit or sheep, is immunized using the isolated protein or protein fragment. The vertebrate may optionally (and preferably) be immunized at least one additional time with the isolated protein or protein fragment, so that the vertebrate exhibits a robust immune response to the protein or protein fragment. Splenocytes are isolated from the immunized vertebrate and fused with an immortalized cell line to form hybridomas, using any of a variety of methods well known in the art. Hybridomas formed in this manner are then screened using standard methods to identify one or more hybridomas which produce an antibody which specifically binds with the protein or protein fragment. The invention also includes hybridomas made by this method and antibodies made using such hybridomas.

[0051] The invention provides methods of diagnosing MS, or determining the risk of developing MS. These methods involve isolating a sample from a subject (e.g., a sample containing blood cells or brain cells), detecting the presence, quantity and/or activity of one

or more markers of the invention in the sample relative to a second sample from a non-diseased subject, or from a non-involved tissue in the same subject. The levels of markers in the two samples are compared, and a substantial increase or decrease in one or more markers in the test sample indicates the presence or risk of presence of MS in the subject.

[0052] The invention also provides methods of assessing the efficacy of a test compound or therapy for inhibiting MS in a subject. These methods involve isolating samples from a subject suffering from MS who is undergoing treatment or therapy, and detecting the presence, quantity, and/or activity of one or more markers of the invention in the first sample relative to a second sample. Where a test compound is administered, the first and second samples are preferably sub-portions of a single sample taken from the patient, wherein the first portion is exposed to the test compound and the second portion is not. In one aspect of this embodiment, the substantially different level of expression is a substantially lower level of expression in the first sample, relative to the second. Most preferably, the level of expression in the first sample approximates (i.e., less than a two fold difference from a control) the level of expression in a third control sample, taken from either a non-diseased subject or non-involved tissue.

[0053] Where the efficacy of a therapy is being assessed, the first sample obtained from the subject is preferably obtained prior to provision of at least a portion of the therapy, whereas the second sample is obtained following provision of the portion of the therapy. The levels of markers in the samples are

compared, preferably against a third control sample as well, and correlated with the presence, risk of presence, or severity of MS. Most preferably, the level of markers in the second sample approximates the level of expression of a third control sample. By assessing whether expression of MS has been lessened or alleviated in the sample, the ability of the treatment or therapy to treat MS is determined.

[0054] The invention also provides a method of screening test compounds for inhibitors of MS, and to the pharmaceutical compositions comprising the test compounds. The method of screening comprises obtaining samples of diseased or involved cells, maintaining separate aliquots of the samples with a plurality of test compounds, and comparing expression of a marker in each of the aliquots to determine whether any of the test compounds provides a substantially different level of expression from a control. In addition, methods of screening may be devised by combining a test compound with a protein and thereby determining the effect of the test compound on the protein. Alternatively, the invention is further directed to a method of screening for bioactive agents capable of interfering with the binding of a protein encoded by the markers of Tables 1-5 (murine) or 7-10 (human), and an antibody, by combining the bioactive agent, protein, and antibody together and determining whether binding of the antibody and protein occurs.

[0055] Moreover, the invention is directed to pharmaceutical compositions comprising the test compound, or bioactive agent, which may further include a marker protein and/or nucleic acid of the invention (e.g., for those markers in Tables 1-5 or 7-10 which

are decreased or increased in quantity or activity in MS versus non-diseased tissue), and can be formulated as described herein. Alternatively, these compositions may include an antibody which specifically binds to a marker protein of the invention and/or an antisense nucleic acid molecule which is complementary to a marker nucleic acid of the invention (e.g., for those markers which are increased in quantity in MS tissue) and can be formulated as described herein.

[0056] The invention further provides methods of modulating a level of expression of a marker of the invention, comprising administration to the diseased cells of the subject a variety of compositions which correspond to the markers of Tables 1-5 (murine) or 7-10 (human), including proteins or antisense oligonucleotides. The protein may be provided to the diseased cells by further providing a vector comprising a polynucleotide encoding the protein to the cells. Alternatively, the expression levels of the markers of the invention may be modulated by providing an antibody, a plurality of antibodies or an antibody conjugated to a therapeutic moiety. Treatment with the antibody may further be localized to the diseased tissue. In another aspect, the invention provides methods for localizing a therapeutic moiety to diseased tissue comprising exposing the tissue to an antibody which is specific to a protein encoded from the markers of the invention. This method may therefore provide a means to inhibit or enhance expression of a specific gene corresponding to a marker listed in Tables 1-5 or 7-10. Where the gene is up-regulated as a result of MS pathology, it is likely that inhibition of MS progression would involve inhibiting expression of the up-regulated gene. As a corollary to this method,

where the gene is down-regulated, inhibition of MS progression would therefore likely require enhancing expression of the down-regulated gene.

[0057] In another aspect, the invention includes antibodies that are specific to proteins corresponding to markers of the invention. Preferably the antibodies are monoclonal, and most preferably, the antibodies are humanized, as per the description of antibodies described below.

[0058] In still another aspect of the invention, the invention includes peptides or proteins which are encoded from the markers of the invention, and to compositions thereof.

[0059] The invention also provides kits for diagnosing a subject with MS, the kit comprising reagents for assessing expression of the markers of the invention. Preferably, the reagents may be an antibody or fragment thereof, wherein the antibody or fragment thereof specifically binds with a protein corresponding to a marker from Tables 1-5 or 7-10. Optionally, the kits may comprise a nucleic acid probe wherein the probe specifically binds with a transcribed polynucleotide corresponding to a marker selected from the group consisting of the markers listed in Tables 1-5 or 7-10.

[0060] The invention further provides kits for assessing the suitability of each of a plurality of compounds for inhibiting progression of MS in a subject. Such kits include a plurality of compounds to be tested, and a reagent (i.e. antibody specific to

corresponding proteins of the invention) for assessing expression of a marker listed in Tables 1-5 or 7-10.

[0061] Modifications to the above-described compositions and methods of the invention, according to standard techniques, will be readily apparent to one skilled in the art and are meant to be encompassed by the invention.

[0062] To facilitate an understanding of the present invention, a number of terms and phrases are defined below:

[0063] As used herein, the term "modulation" includes, in its various grammatical forms (e.g., "modulated", "modulation", "modulating", etc.), up-regulation, induction, stimulation, potentiation, and/or relief of inhibition, as well as inhibition and/or down-regulation.

[0064] As used herein, the terms "polynucleotide" and "oligonucleotide" are used interchangeably, and include polymeric forms of nucleotides of any length, either deoxyribonucleotides or ribonucleotides, or analogs thereof. Polynucleotides may have any three-dimensional structure, and may perform any function, known or unknown. The following are non-limiting examples of polynucleotides: a gene or gene fragment, exons, introns, messenger RNA (mRNA), transfer RNA, ribosomal RNA, ribozymes, cDNA, recombinant polynucleotides, branched polynucleotides, plasmids, vectors, isolated DNA of any sequence, isolated RNA of any sequence, nucleic acid probes, and primers. A polynucleotide may comprise modified nucleotides, such

as methylated nucleotides and nucleotide analogs. If present, modifications to the nucleotide structure may be imparted before or after assembly of the polymer. The sequence of nucleotides may be interrupted by non-nucleotide components. A polynucleotide may be further modified after polymerization, such as by conjugation with a labeling component. The term also includes both double- and single-stranded molecules. Unless otherwise specified or required, any embodiment of this invention that is a polynucleotide encompasses both the double-stranded form and each of two complementary single-stranded forms known or predicted to make up the double-stranded form.

[0065] A polynucleotide is composed of a specific sequence of four nucleotide bases: adenine (A); cytosine(C); guanine (G); thymine (T); and uracil (U) for guanine when the polynucleotide is RNA. This, the term "polynucleotide sequence" is the alphabetical representation of a polynucleotide molecule. This alphabetical representation can be inputted into databases in a computer having a central processing unit and used for bioinformatics applications such as functional genomics and homology searching.

[0066] A "gene" includes a polynucleotide containing at least one open reading frame that is capable of encoding a particular polypeptide or protein after being transcribed and translated. Any of the polynucleotide sequences described herein may be used to identify larger fragments or full-length coding sequences of the gene with which they are associated. Methods of isolating larger fragment sequences are known to those of sill in the art, some of which are described herein.

[0067] A "gene product" includes an amino acid sequence(*e.g.*, peptide or polypeptide) generated when a gene is transcribed and translated.

[0068] As used herein, a "polynucleotide corresponds to" another (a first) polynucleotide if it is related to the first polynucleotide by any of the following relationships:

1) The second polynucleotide comprises the first polynucleotide and the second polynucleotide encodes a gene product.

2) The second polynucleotide is 5' or 3' to the first polynucleotide in cDNA, RNA, genomic DNA, or fragments of any of these polynucleotides. For example, a second polynucleotide may be a fragment of a gene that includes the first and second polynucleotides. The first and second polynucleotides are related in that they are components of the gene coding for a gene product, such as a protein or antibody. However, it is not necessary that the second polynucleotide comprises or overlaps with the first polynucleotide to be encompassed within the definition of "corresponding to" as used herein. For example, the first polynucleotide may be a fragment of a 3' untranslated region of the second polynucleotide. The first and second polynucleotide may be fragments of a gene coding for a gene product. The second polynucleotide may be an exon of the gene while the first polynucleotide may be an intron of the gene.

3) The second polynucleotide is the complement of the first polynucleotide.

[0069] As used herein, the term, "transcribed" or "transcription" refers to the process by which genetic code information is transferred from one kind of nucleic acid to another, and refers in particular to the process by which a base sequence of mRNA is synthesized on a template of cDNA.

[0070] A "probe" when used in the context of polynucleotide manipulation includes an oligonucleotide that is provided as a reagent to detect a target present in a sample of interest by hybridizing with the target. Usually, a probe will comprise a label or a means by which a label can be attached, either before or subsequent to the hybridization reaction. Suitable labels include, but are not limited to radioisotopes, fluorochromes, chemiluminescent compounds, dyes, and proteins, including enzymes.

[0071] A "primer" includes a short polynucleotide, generally with a free 3'-OH group that binds to a target or "template" present in a sample of interest by hybridizing with the target, and thereafter promoting polymerization of a polynucleotide complementary to the target. A "polymerase chain reaction" ("PCR") is a reaction in which replicate copies are made of a target polynucleotide using a "pair of primers" or "set or primers" consisting of "upstream" and a "downstream" primer, and a catalyst of polymerization, such as a DNA polymerase, and typically a thermally-stable polymerase enzyme. Methods for PCR are well known in the art, and are taught, for example, in MacPherson et al., IRL Press at Oxford University Press (1991)). All

processes of producing replicate copies of a polynucleotide, such as PCR or gene cloning, are collectively referred to herein as "replication". A primer can also be used as a probe in hybridization reactions, such as Southern or Northern blot analyses (see, e.g., Sambrook, J., Fritsh, E.F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

[0072] The term "cDNAs" includes complementary DNA, that is mRNA molecules present in a cell or organism made into cDNA with an enzyme such as reverse transcriptase. A "cDNA library" includes a collection of mRNA molecules present in a cell or organism, converted into cDNA molecules with the enzyme reverse transcriptase, then inserted into "vectors" (other DNA molecules that can continue to replicate after addition of foreign DNA). Exemplary vectors for libraries include bacteriophage, viruses that infect bacteria (e.g., lambda phage). The library can then be probed for the specific cDNA (and thus mRNA) of interest.

[0073] A "gene delivery vehicle" includes a molecule that is capable of inserting one or more polynucleotides into a host cell. Examples of gene delivery vehicles are liposomes, biocompatible polymers, including natural polymers and synthetic polymers; lipoproteins; polypeptides; polysaccharides; lipopolysaccharides; artificial viral envelopes; metal particles; and bacteria, viruses and viral vectors, such as baculovirus, adenovirus, and retrovirus, bacteriophage, cosmid, plasmid, fungal vector and other recombination vehicles typically used in the art which

have been described for replication and/or expression in a variety of eukaryotic and prokaryotic hosts. The gene delivery vehicles may be used for replication of the inserted polynucleotide, gene therapy as well as for simply polypeptide and protein expression.

[0074] A "vector" includes a self-replicating nucleic acid molecule that transfers an inserted polynucleotide into and/or between host cells. The term is intended to include vectors that function primarily for insertion of a nucleic acid molecule into a cell, replication vectors that function primarily for the replication of nucleic acid and expression vectors that function for transcription and/or translation of the DNA or RNA. Also intended are vectors that provide more than one of the above function.

[0075] A "host cell" is intended to include any individual cell or cell culture which can be or has been a recipient for vectors or for the incorporation of exogenous nucleic acid molecules, polynucleotides and/or proteins. It also is intended to include progeny of a single cell. The progeny may not necessarily be completely identical (in morphology or in genomic or total DNA complement) to the original parent cell due to natural, accidental, or deliberate mutation. The cells may be prokaryotic or eukaryotic, and include but are not limited to bacterial cells, yeast cells, insect cells, animal cells, and mammalian cells, *e.g.*, murine, rat, simian or human cells.

[0076] The term "genetically modified" includes a cell containing and/or expressing a foreign gene or nucleic acid sequence which in turn modifies the genotype or

phenotype of the cell or its progeny. This term includes any addition, deletion, or disruption to a cell's endogenous nucleotides.

[0077] As used herein, "expression" includes the process by which polynucleotides are transcribed into mRNA and translated into peptides, polypeptides, or proteins. If the polynucleotide is derived from genomic DNA, expression may include splicing of the mRNA, if an appropriate eukaryotic host is selected. Regulatory elements required for expression include promoter sequences to bind RNA polymerase and transcription initiation sequences for ribosome binding. For example, a bacterial expression vector includes a promoter such as the lac promoter and for transcription initiation the Shine-Dalgarno sequence and the start codon AUG (Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989). Similarly, a eukaryotic expression vector includes a heterologous or homologous promoter for RNA polymerase II, a downstream polyadenylation signal, the start codon AUG, and a termination codon for detachment of the ribosome. Such vectors can be obtained commercially or assembled by the sequences described in methods well known in the art, for example, the methods described below for constructing vectors in general.

[0078] "Differentially expressed", as applied to a gene, includes the differential production of mRNA transcribed from a gene or a protein product encoded by the gene. A differentially expressed gene may be overexpressed or underexpressed as compared to the expression level of a normal or control cell. In one

aspect, it includes a differential that is 2 times, preferably 5 times or preferably 10 times higher or lower than the expression level detected in a control sample. The term "differentially expressed" also includes nucleotide sequences in a cell or tissue which are expressed where silent in a control cell or not expressed where expressed in a control cell.

[0079] The term "polypeptide" includes a compound of two or more subunit amino acids, amino acid analogs, or peptidomimetics. The subunits may be linked by peptide bonds. In another embodiment, the subunit may be linked by other bonds, *e.g.*, ester, ether, etc. As used herein the term "amino acid" includes either natural and/or unnatural or synthetic amino acids, including glycine and both the D or L optical isomers, and amino acid analogs and peptidomimetics. A peptide of three or more amino acids is commonly referred to as an oligopeptide. Peptide chains of greater than three or more amino acids are referred to as a polypeptide or a protein.

[0080] "Hybridization" includes a reaction in which one or more polynucleotides react to form a complex that is stabilized via hydrogen bonding between the bases of the nucleotide residues. The hydrogen bonding may occur by Watson-Crick base pairing, Hoogstein binding, or in any other sequence-specific manner. The complex may comprise two strands forming a duplex structure, there or more strands forming a multi-stranded complex, a single self-hybridizing strand, or any combination of these. A hybridization reaction may constitute a step in a more extensive process, such as the initiation of a PCR reaction, or the enzymatic cleavage of a polynucleotide by a ribozyme.

[0081] Hybridization reactions can be performed under conditions of different "stringency". The stringency of a hybridization reaction includes the difficulty with which any two nucleic acid molecules will hybridize to one another. The present invention also includes polynucleotides capable of hybridizing under reduced stringency conditions, more preferably stringent conditions, and most preferably highly stringent conditions, to polynucleotides described herein. Examples of stringency conditions are shown in Table A below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

Table A. Stringency Conditions

Stringency Condition	Poly-nucleotide Hybrid	Hybrid Length (bp)[1]	Hybridization Temperature and Buffer[H]	Wash Temperature and Buffer[H]
A	DNA:DNA	> 50	65°C; 1xSSC -or- 42°C; 1xSSC, 50% formamide	65°C; 0.3xSSC
B	DNA:DNA	<50	T_B*; 1xSSC	T_B*; 1xSSC
C	DNA:RNA	> 50	67°C; 1xSSC -or- 45°C; 1xSSC, 50% formamide	67°C; 0.3xSSC
D	DNA:RNA	<50	T_D*; 1xSSC	T_D*; 1xSSC
E	RNA:RNA	>50	70°C; 1xSSC -or- 50°C; 1xSSC, 50% formamide	70°C; 0.3xSSC
F	RNA:RNA	<50	T_F*; 1xSSC	T_F*; 1xSSC
G	DNA:DNA	> 50	65°C; 4xSSC -or- 42°C; 4xSSC, 50% formamide	65°C; 1xSSC
H	DNA:DNA	<50	T_H*; 4xSSC	T_H*; 4xSSC
I	DNA:RNA	> 50	67°C; 4xSSC -or- 45°C; 4xSSC, 50% formamide	67°C; 1xSSC
J	DNA:RNA	<50	T_J*; 4xSSC	T_J*; 4xSSC
K	RNA:RNA	> 50	70°C; 4xSSC -or- 50°C; 4xSSC, 50% formamide	67°C; 1xSSC
L	RNA:RNA	<50	T_L*; 2xSSC	T_L*; 2xSSC
M	DNA:DNA	> 50	50°C; 4xSSC -or- 40°C; 6xSSC, 50% formamide	50°C; 2xSSC
N	DNA:DNA	<50	T_N*; 6xSSC	T_N*; 6xSSC
O	DNA:RNA	> 50	55°C; 4xSSC -or- 42°C; 6xSSC, 50% formamide	55°C; 2xSSC
P	DNA:RNA	<50	T_P*; 6xSSC	T_P*; 6xSSC
Q	RNA:RNA	> 50	60°C; 4xSSC -or- 45°C; 6xSSC, 50% formamide	60°C; 2xSSC
R	RNA:RNA	<50	T_R*; 4xSSC	T_R*; 4xSSC

1: The hybrid length is that anticipated for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide to a target polynucleotide of unknown sequence, the hybrid length is assumed to be that of the hybridizing polynucleotide. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarity.

[H]: SSPE (1xSSPE is 0.15M NaCl, 10mM NaH_2PO_4, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete.

T_B* - T_R*: The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length, T_m(°C) = 2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, T_m(°C) = 81.5 + 16.6($\log_{10}Na^+$) + 0.41(%G+C) - (600/N), where N is the number of bases in the hybrid, and Na^+ is the concentration of sodium ions in the hybridization buffer (Na^+ for 1xSSC = 0.165 M).

Additional examples of stringency conditions for polynucleotide hybridization are provided in Sambrook, J., E.F. Fritsch, and T. Maniatis, 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, chapters 9 and 11, and Current Protocols in Molecular Biology, 1995, F.M. Ausubel et al., eds., John Wiley & Sons, Inc., sections 2.10 and 6.3-6.4, incorporated herein by reference.

[0082] When hybridization occurs in an antiparallel configuration between two single-stranded polynucleotides, the reaction is called "annealing" and those polynucleotides are described as "complementary". A double-stranded polynucleotide can be "complementary" or "homologous" to another polynucleotide, if hybridization can occur between one of the strands of the first polynucleotide and the second. "Complementarity" or "homology" (the degree that one polynucleotide is complementary with another) is quantifiable in terms of the proportion of bases in opposing strands that are expected to hydrogen bond with each other, according to generally accepted base-pairing rules.

[0083] An "antibody" includes an immunoglobulin molecule capable of binding an epitope present on an antigen. As used herein, the term encompasses not only intact immunoglobulin molecules such as monoclonal and polyclonal antibodies, but also anti-idotypic antibodies, mutants, fragments, fusion proteins, bi-specific antibodies, humanized proteins, and modifications of the immunoglobulin molecule that comprises an antigen recognition site of the required specificity.

[0084] As used herein, the term "diseased" refers to cells, tissues or samples from a subject afflicted with multiple sclerosis, wherein the cell, tissue or sample has been affected by multiple sclerosis (i.e. from white-matter lesion). As used herein, the term "non-diseased" refers to cells, tissues or other such samples taken from a subject who is not afflicted with multiple sclerosis. As used herein, "non-involved" refers to cells, tissues, or samples wherein the tissue is from a subjected afflicted with MS, but wherein the cells, tissues or samples are believed to be unaffected by multiple sclerosis. Preferred tissue (and cell) samples are from brain, blood, sera, lymph, thymus, spleen, bone marrow or pus. Most preferred samples are peripheral blood mononuclear cells ("PBMC") or brain tissue.

[0085] As used herein, the term "marker" includes a polynucleotide or polypeptide molecule which is present or absent, or increased or decreased in quantity or activity in subjects afflicted with multiple sclerosis, or in MS-associated cells. The relative change in quantity or activity of the marker is correlated with the incidence or risk of incidence of multiple sclerosis.

[0086] As used herein, the term "panel of markers" includes a group of markers, the quantity or activity of each member of which is correlated with the incidence or risk of incidence of a MS-associated condition. In certain embodiments, a panel of markers may include only those markers which are either increased or decreased in quantity or activity in subjects afflicted with or cells involved in a MS-associated condition. In a preferred embodiment, the

panel of markers comprises at least 5 markers, and most preferably, the panel comprises markers listed in Table 10. In other embodiments, a panel of markers may include only those markers present in a specific tissue type which are correlated with the incidence of risk of incidence of a MS-associated condition.

Various aspects of the invention are described in further detail in the following subsections:

I. Isolated Nucleic Acid Molecules

[0087] One aspect of the invention pertains to isolated nucleic acid molecules that either themselves are the genetic markers (*e.g.*, mRNA) of the invention, or which encode the polypeptide markers of the invention, or fragments thereof. Another aspect of the invention pertains to isolated nucleic acid fragments sufficient for sue as hybridization probes to identify the nucleic acid molecules encoding the markers for the invention in a sample, as well as nucleotide fragments for use as PCR primers of the amplification or mutation of the nucleic acid molecules which encode the markers of the invention. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (*e.g.*, cDNA or genomic DNA) and RNA molecules (*e.g.*, mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA..

[0088] The term "isolated nucleic acid molecule" includes nucleic acid molecules which are separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. For example,

with regards to genomic DNA, the term "isolated" includes nucleic acid molecules which are separated from the chromosome with which the genomic DNA is naturally associated. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated marker nucleic acid molecule of the invention, or nucleic acid molecule encoding a polypeptide marker of the invention, can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized.

[0089] A nucleic acid molecule of the present invention, *e.g.*, a nucleic acid molecule having the nucleotide sequence of one of the genes set forth in Tables 1-5 or 7-10, or a portion thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or portion of the nucleic acid sequence of one of the genes set forth in Tables 1-5 or 7-10 as a hybridization probe, a marker gene of the invention or a nucleic acid molecule encoding a polypeptide marker of the invention can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook, J., Fritsh, E. F., and Maniatis,

T. *Molecular Cloning: A Laboratory Manual 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold spring Harbor, NY, 1989).

[0090] A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to marker nucleotide sequences, or nucleotide sequences encoding a marker of the invention can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

[0091] In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a complement of the nucleotide sequence of a marker of the invention (*e.g.*, a gene set forth in Tables 1-5 or 7-10), or a portion of any of these nucleotide sequences. A nucleic acid molecule which is complementary to such a nucleotide sequence is one which is sufficiently complementary t the nucleotide sequence such that it can hybridize to the nucleotide sequence, thereby forming a stable duplex.

[0092] The nucleic acid molecule of the invention, moreover, can comprise only a portion of the nucleic acid sequence of a marker nucleic acid of the invention, or a gene encoding a marker polypeptide of the invention, for example, a fragment which can be used as a probe or primer. The probe/primer typically

comprises substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 7 or 15, preferably about 20 or 25, more preferably about 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 400 or more consecutive nucleotides of a marker nucleic acid, or a nucleic acid encoding a marker polypeptide of the invention.

[0093] Probes based on the nucleotide sequence of a marker gene or of a nucleic acid molecule encoding a marker polypeptide of the invention can be used to detect transcripts or genomic sequences corresponding to the marker gene(s) and/or marker polypeptide(s) of the invention. In preferred embodiments, the probe comprises a label group attached thereto, *e.g.*, the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells or tissue which misexpress (*e.g.*, over- or under-express) a marker polypeptide of the invention, or which have greater or fewer copies of a marker gene of the invention. For example, a level of a marker polypeptide-encoding nucleic acid in a sample of cells from a subject may be detected, the amount of mRNA transcript of a gene encoding a marker polypeptide may be determined, or the presence of mutations or deletions of a marker gene of the invention may be assessed.

[0094] The invention further encompasses nucleic acid molecules that differ from the nucleic acid sequences of the genes set forth in Tables 1-5 or 7-10, due to degeneracy of the genetic code and which thus encode

the same proteins as those encoded by the genes shown in Tables 1-5 or 7-10.

[0095] In addition to the nucleotide sequences of the genes set forth in Tables 1-5 or 7-10, it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of the proteins encoded by the genes set forth in Tables 1-5 or 7-10 may exist within a population *e.g.*, the human population). Such genetic polymorphism in the genes set forth in Tables 1-5 or 7-10 may exist among individuals within a population due to natural allelic variation. An allele is one of a group of genes which occur alternatively at a given genetic locus. In addition it will be appreciated that DNA polymorphisms that affect RNA expression levels can also exist that may affect the overall expression level of that gene *e.g.*, by affecting regulation or degradation). As used herein, the phrase "allelic variant" includes a nucleotide sequence which occurs ta a given locus or to a polypeptide encoded by the nucleotide sequence. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules which include an open reading frame encoding a marker polypeptide of the invention.

[0096] Nucleic acid molecules corresponding to natural allelic variants and homologues of the marker genes, or genes encoding the marker proteins of the invention can be isolated based on their homology to the genes set forth in Tables 7-10, using the cDNAs disclosed herein, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions. Nucleic acid molecules corresponding to natural allelic variants and

homologues of the marker genes of the invention can further be isolated by mapping to the same chromosome or locus as the marker genes or genes encoding the marker proteins of the invention.

[0097] In another embodiment, an isolated nucleic acid molecule of the invention is at least 15, 20, 25, 30, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000 or more nucleotides in length and hybridizes under stringent conditions to a nucleic acid molecule corresponding to a nucleotide sequence of a marker gene or gene encoding a marker protein of the invention. As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 60% homologous to each other typically remain hybridized to each other. Preferably, the conditions are such that sequences at least about 70%, more preferably at least about 80%, even more preferably at least about 85% or 90% homologous to each other typically remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. Preferably, an isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequence of one of the genes set forth in Tables 1-5 or 7-10 corresponds to a naturally-occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule includes an RNA or DNA molecule having a nucleotide sequence that occurs in nature (*e.g.*, encodes a natural protein).

[0098] In addition to naturally-occurring allelic variants of the marker gene and gene encoding a marker protein of the invention sequences that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation into the nucleotide sequences of the marker genes or genes encoding the marker proteins of the invention, thereby leading to changes in the amino acid sequence of the encoded proteins, without altering the functional activity of these proteins. For example, nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues can be made. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of a protein without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For example, amino acid residues that are conserved among allelic variants or homologs of a gene (*e.g.*, among homologs of a gene from different species) are predicted to be particularly unamenable to alteration.

[0099] Accordingly, another aspect of the invention pertains to nucleic acid molecules encoding a marker protein of the invention that contain changes in amino acid residues that are not essential for activity. Such proteins differ in amino acid sequence from the marker proteins encoded by the genes set forth in Tables 1-5 or 7-10, yet retain biological activity. In one embodiment, the protein comprises an amino acid sequence at least about 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or more homologous to a marker protein of the invention.

[0100] An isolated nucleic acid molecule encoding a protein homologous to a marker protein of the invention can be created by introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of the gene encoding the marker protein, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein. Mutations can be introduced into the genes of the invention (*e.g.*, a gene set forth in Tables 7-10) by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Alternatively, mutations can be introduced randomly along all or part of a coding sequence of a gene of the invention, such as by saturation mutagenesis, and the resultant mutants can be screened for biological activity to identify mutants that retain activity. Following mutagenesis, the

encoded protein can be expressed recombinantly and the activity of the protein can be determined.

[0101] Another aspect of the invention pertains to isolated nucleic acid molecules which are antisense to the marker genes and genes encoding marker proteins of the invention. An "antisense" nucleic acid comprises a nucleotide sequence which is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. Accordingly, an antisense nucleic acid can hydrogen bond to a sense nucleic acid. The antisense nucleic acid can be complementary to an entire coding strand of a gene of the invention (*e.g.*, a gene set forth in Tables 1-5 or 7-10), or to only a portion thereof. In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence of the invention. The term "coding region" includes the region of the nucleotide sequence comprising codons which are translated into amino acid. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence of the invention.

[0102] The term "noncoding region" includes 5' and 3' sequences which flank the coding region that are not translated into amino acids (i.e., also referred to as 5' and 3' untranslated regions).

[0103] Antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick base pairing. The antisense nucleic acid molecule can

be complementary to the entire coding region of an mRNA corresponding to a gene of the invention, but more preferably is an oligonucleotide which is antisense to only a portion of the coding or noncoding region. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioatc derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine, 5-(carboxyhydroxyhnethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladen4exine, unacil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5- oxyacetic acid

methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

[0104] The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a marker protein of the invention to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the cases of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention include direct injection at a tissue site (*e.g.*, in brain). Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the

vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

[0105] In yet another embodiment, the antisense nucleic acid molecule of the invention is an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids. Res.* 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res.* 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue *et al.* (1987) *FEBS Lett.* 215:327-330).

[0106] In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoif and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave mRNA transcripts of the genes of the invention (*e.g.*, a gene set forth in Tables 1-5 or 7-10) to thereby inhibit translation of this mRNA. A ribozyme having specificity for a marker protein-encoding nucleic acid can be designed based upon the nucleotide sequence of a gene of the invention, disclosed herein. For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which

the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a marker protein-encoding mRNA. See, *e.g.*, Cech *et al.* U.S. Patent No. 4,987,071; and Cech *et al.* U.S. Patent No. 5,116,742. Alternatively, mRNA transcribed from a gene of the invention can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel, D. and Szostak, J.W. (1993) *Science* 261:1411-1418.

[0107] Alternatively, expression of a gene of the invention (*e.g.*, a gene set forth in Tables 1-5 or 7-10) can be inhibited by targeting nucleotide sequences complementary to the regulatory region of these genes (*e.g.*, the promoter and/or enhancers) to form triple helical structures that prevent transcription of the gene in target cells. See generally, Helene, C. (1991) *Anticancer Drug Des.* 6(6):569-84; Helene, C. *et al.* (1992) *Ann. N. Y. Acad Sci.* 660:27-36; and Maher, L.J. (1992) *Bioassays* 14(12):807-15.

[0108] Expression of the marker genes, and genes encoding marker proteins of the invention, can also be inhibited using RNA interference ("RNA_i"). This is a technique for post transcriptional gene silencing ("PTGS"), in which target gene activity is specifically abolished with cognate double-stranded RNA ("dsRNA"). RNA_i resembles in many aspects PTGS in plants and has been detected in many invertebrates including trypanosome, hydra, planaria, nematode and fruit fly (Drosophila melanogaster). It may be involved in the modulation of transposable element mobilization and antiviral state formation. RNA_i in mammalian systems is disclosed in PCT application WO 00/63364 which is

incorporated by reference herein in its entirety. Basically, dsRNA of at least about 600 nucleotides, homologous to the target marker is introduced into the cell and a sequence specific reduction in gene activity is observed. See generally, Ui-Teia, K. et al. *FEBS Letters* 479: 79-82.

[0109] In yet another embodiment, the nucleic acid molecules of the present invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acid molecules can be modified to generate peptide nucleic acids (see Hyrup B. *et al.* (1996) *Bioorganic & Medicinal Chemistry* 4(1): 5 23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup B. *et al.* (1996) *supra;* Perry-O'Keefe *et al.* Proc. Natl. Acad. Sci. 93: 14670-675.

[0110] PNAs can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, for example, inducing transcription or translation arrest or inhibiting replication. PNAs of the nucleic acid molecules of the

invention (*e.g.*, a gene set forth in Tables 1-5 or 7-10) can also be used in the analysis of single base pair mutations in a gene, (*e.g.*, by PNA-directed PCR clamping); as 'artificial restriction enzymes' when used in combination with other enzymes, (*e.g.*, S1 nucleases (Hyrup B. (1996) *supra*)); or as probes or primers for DNA sequencing or hybridization (Hyrup B. *et al.* (1996) *supra*; Perry-O'Keefe *supra*).

[0111] In another embodiment, PNAs can be modified, (*e.g.*, to enhance their stability or cellular uptake), by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of the nucleic acid molecules of the invention can be generated which may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, (*e.g.*, RNAse H and DNA polymerases), to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup B. (1996) *supra*). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup B. (1996) *supra* and Finn P.J. *et al.* (1996) *Nucleic Acids Res.* 24 (17): 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used as a between the PNA and the 5' end of DNA (Mag, M. *et al.* (1989) *Nucleic Acid Res.* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to

produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn P.J. *et al.* (1996) *supra*). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment (Peterser, K.H. *et al.* (1975) *Bioorganic Med Chem. Lett.* 5: 1119-11124).

[0112] In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:6553-6556; Lemaitre *et al.* (1987) *Pros. Natl. Acad Sci. USA* 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization-triggered cleavage agents (See, *e.g.*, Krol *et al.* (1988) *Bio-Techniques* 6:958-976) or intercalating agents. (See, *e.g.*, Zon (1988) *Pharm. Res.* 5:539-549). To this end, the oligonucleotide may be conjugated to another molecule, (*e.g.*, a peptide, hybridization triggered cross-linking agent, transport agent, or hybridization-triggered cleavage agent). Finally, the oligonucleotide may be detectably labeled, either such that the label is detected by the addition of another reagent (*e.g.*, a substrate for an enzymatic label), or is detectable immediately upon hybridization of the nucleotide (*e.g.*, a radioactive label or a fluorescent label (*e.g.*, a molecular beacon, as described in U.S. Patent 5,876,930).

II. Isolated Proteins and Antibodies

[0113] One aspect of the invention pertains to isolated marker proteins, and biologically active portions thereof, as well as polypeptide fragments suitable for use as immunogens to raise anti-marker protein antibodies. In one embodiment, native marker proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, marker proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, a marker protein or polypeptide can be synthesized chemically using standard peptide synthesis techniques.

[0114] An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the marker protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of marker protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of marker protein having less than about 30% (by dry weight) of non-marker protein (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-marker protein, still more preferably less than about 10% of non-marker protein, and most preferably less than about 5% non-marker protein. When the marker protein or biologically active portion thereof is recombinantly

produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation.

[0115] The language "substantially free of chemical precursors or other chemicals" includes preparations of marker protein in which the protein is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of protein having less than about 30% (by dry weight) of chemical precursors or non-protein chemicals, more preferably less than about 20% chemical precursors or non-protein chemicals, still more preferably less than about 10% chemical precursors or non-protein chemicals, and most preferably less than about 5% chemical precursors or non-protein chemicals.

[0116] As used herein, a "biologically active portion" of a marker protein includes a fragment of a marker protein comprising amino acid sequences sufficiently homologous to or derived from the amino acid sequence of the marker protein, which include fewer amino acids than the full length marker proteins, and exhibit at least one activity of a marker protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the marker protein. A biologically active portion of a marker protein can be a polypeptide which is, for example, 10, 25, 50, 100, 200 or more amino acids in length. Biologically active portions of a marker protein can be used as targets for

developing agents which modulate a marker protein-mediated activity.

[0117] In a preferred embodiment, marker protein is encoded by a gene set forth in Tables 1-5 or 7-10. In other embodiments, the marker protein is substantially homologous to a marker protein encoded by a gene set forth in Tables 1-5 or 7-10, and retains the functional activity of the marker protein, yet differs in amino acid sequence due to natural allelic variation or mutagenesis, as described in detail in subsection I above. Accordingly, in another embodiment, the marker protein is a protein which comprises an amino acid sequence at least about 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or more homologous to the amino acid sequence encoded by a gene set forth in Tables 1-5 or 7-10.

[0118] To determine the percent identity of two amino acid sequences or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). In a preferred embodiment, the length of a reference sequence aligned for comparison purposes is at least 30%, preferably at least 40%, more preferably at least 50%, even more preferably at least 60%, and even more preferably at least 70%, 80%, or 90% of the length of the reference sequence. The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in

the second sequence, then the molecules are identical at that position (as used herein amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid "homology"). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences.

[0119] The comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (*J. Mot. Biol.* (48):444-453 (1970)) algorithm which has been incorporated into the GAP program in the GCG software package (available at http://www.gcg.com), using either a Blossom 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another preferred embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package (available at http://www.gcg.com), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. In another embodiment, the percent identity between two amino acid or nucleotide sequences is determined using the algorithm of E. Meyers and W. Miller (CABIOS, 4:11-17 (1989)) which has been incorporated into the ALIGN program (version 2.0), using a PAM 120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

[0120] The nucleic acid and protein sequences of the present invention can further be used as a "query sequence" to perform a search against public databases to, for example, identify other family members or related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, *et al.* (1990) *J. Mol. Biol.* 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to marker protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.*, (1997) *Nucleic Acids Res.* 25(17):3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (*e.g.*, XBLAST and NBLAST) can be used. See http://www.ncbi.nim.nih.gov.

[0121] The invention also provides chimeric or fusion marker proteins. As used herein, a marker "chimeric protein" or "fusion protein" comprises a marker polypeptide operatively linked to a non-marker polypeptide. An "marker polypeptide" includes a polypeptide having an amino acid sequence encoded by a gene set forth in Tables 1-5 or 7-10, whereas a "non-marker polypeptide" includes a polypeptide having an amino acid sequence corresponding to a protein which is not substantially homologous to the marker protein, *e.g.*, a protein which is different from marker protein and which is derived from the same or a different

organism. Within a marker fusion protein the polypeptide can correspond to all or a portion of a marker protein. In a preferred embodiment, a marker fusion protein comprises at least one biologically active portion of a marker protein. Within the fusion protein, the term "operatively linked" is intended to indicate that the marker polypeptide and the non-marker polypeptide are fused in-frame to each other. The non-marker polypeptide can be fused to the N-terminus or C-terminus of the marker polypeptide.

[0122] For example, in one embodiment, the fusion protein is a GST-marker fusion protein in which the marker sequences are fused to the C-terminus of the GST sequences. Such fusion proteins can facilitate the purification of recombinant marker proteins.

[0123] In another embodiment, the fusion protein is a marker protein containing a heterologous signal sequence at its N-terminus. In certain host cells (*e.g.*, mammalian host cells), expression and/or secretion of marker proteins can be increased through use of a heterologous signal sequence. Such signal sequences are well known in the art.

[0124] The marker fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject in vivo, as described herein. The marker fusion proteins can be used to affect the bioavailability of a marker protein substrate. Use of marker fusion proteins may be useful therapeutically for the treatment of disorders (*e.g.*, multiple sclerosis) caused by, for example, (i) aberrant modification or mutation of a gene encoding a marker

protein; (ii) mis-regulation of the marker protein-encoding gene; and (iii) aberrant post-translational modification of a marker protein.

[0125] Moreover, the marker-fusion proteins of the invention can be used as immunogens to produce anti-marker protein antibodies in a subject, to purify marker protein ligands and in screening assays to identify molecules which inhibit the interaction of a marker protein with a marker protein substrate.

[0126] Preferably, a marker chimeric or fusion protein of the invention is produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, for example by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, *Current Protocols In Molecular Biology*, eds. Ausubel *et al.* John Wiley & Sons: 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A marker protein-encoding nucleic acid can be cloned into such an expression vector such that

the fusion moiety is linked in-frame to the marker protein.

[0127] A signal sequence can be used to facilitate secretion and isolation of the secreted protein or other proteins of interest. Signal sequences are typically characterized by a core of hydrophobic amino acids which are generally cleaved from the mature protein during secretion in one or more cleavage events. Such signal peptides contain processing sites that allow cleavage of the signal sequence from the mature proteins as they pass through the secretory pathway. Thus, the invention pertains to the described polypeptides having a signal sequence, as well as to polypeptides from which the signal sequence has been proteolytically cleaved (*i.e.*, the cleavage products). In one embodiment, a nucleic acid sequence encoding a signal sequence can be operably linked in an expression vector to a protein of interest, such as a protein which is ordinarily not secreted or is otherwise difficult to isolate. The signal sequence directs secretion of the protein, such as from a eukaryotic host into which the expression vector is transformed, and the signal sequence is subsequently or concurrently cleaved. The protein can then be readily purified from the extracellular medium by art recognized methods.

[0128] Alternatively, the signal sequence can be linked to the protein of interest using a sequence which facilitates purification, such as with a GST domain.

[0129] The present invention also pertains to variants of the marker proteins of the invention which function

as either agonists (mimetics) or as antagonists to the marker proteins. Variants of the marker proteins can be generated by mutagenesis, *e.g.*, discrete point mutation or truncation of a marker protein. An agonist of the marker proteins can retain substantially the same, or a subset, of the biological activities of the naturally occurring form of a marker protein. An antagonist of a marker protein can inhibit one or more of the activities of the naturally occurring form of the marker protein by, for example, competitively modulating an activity of a marker protein. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the biological activities of the naturally occurring forth of the protein has fewer side effects in a subject relative to treatment with the naturally occurring form of the marker protein.

[0130] Variants of a marker protein which function as either marker protein agonists (mimetics) or as marker protein antagonists can be identified by screening combinatorial libraries of mutants, *e.g.*, truncation mutants, of a marker protein for marker protein agonist or antagonist activity. In one embodiment, a variegated library of marker protein variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of marker protein variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential marker protein sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (*e.g.*, for phage display) containing

the set of marker protein sequences therein. There are a variety of methods which can be used to produce libraries of potential marker protein variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential marker protein sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, *e.g.*, Narang, S.A. (1983) *Tetrahedron* 39:3; Itakura *et al.* (1984) *Annu. Rev. Biochem.* 53:323; Itakura *et al.* (1984) *Science* 198:1055; Ike *et al.* (1983) *Nucleic Acid Res.* 11:477).

[0131] In addition, libraries of fragments of a protein coding sequence corresponding to a marker protein of the invention can be used to generate a variegated population of marker protein fragments for screening and subsequent selection of variants of a marker protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of a marker protein coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-

terminal, C-terminal and internal fragments of various sizes of the marker protein.

[0132] Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. The most widely used techniques, which are amenable to high-throughput analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify marker variants (Arkin and Yourvan (1992) *Proc. Natl. Acad. Sci. USA* 89:7811-7815; Delgrave *et al.* (1993) *Protein Engineering* 6(3):327-331).

[0133] An isolated marker protein, or a portion or fragment thereof, can be used as an immunogen to generate antibodies that bind marker proteins using standard techniques for polyclonal and monoclonal antibody preparation. A full-length marker protein can be used or, alternatively, the invention provides antigenic peptide fragments of these proteins for use as immunogens. The antigenic peptide of a marker protein comprises at least 8 amino acid residues of an amino acid sequence encoded by a gene set forth in Tables 1-5 or 7-10, and encompasses an epitope of a

marker protein such that an antibody raised against the peptide forms a specific immune complex with the marker protein. Preferably, the antigenic peptide comprises at least 10 amino acid residues, more preferably at least 15 amino acid residues, even more preferably at least 20 amino acid residues, and most preferably at least 30 amino acid residues.

[0134] Preferred epitopes encompassed by the antigenic peptide are regions of the marker protein that are located on the surface of the protein, *e.g.*, hydrophilic regions, as well as regions with high antigenicity.

[0135] A marker protein immunogen typically is used to prepare antibodies by immunizing a suitable subject, (*e.g.*, rabbit, goat, mouse or other mammal) with the immunogen. An appropriate immunogenic preparation can contain, for example, recombinantly expressed marker protein or a chemically synthesized marker polypeptide. The preparation can further include an adjuvant, such as Freund's complete or incomplete adjuvant, or similar immunostimulatory agent. Immunization of a suitable subject with an immunogenic marker protein preparation induces a polyclonal anti-marker protein antibody response.

[0136] Accordingly, another aspect of the invention pertains to anti-marker protein antibodies. The term "antibody" as used herein includes immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, i.e., molecules that contain an antigen binding site which specifically binds (immunoreacts with) an antigen, such as a marker

protein. Examples of immunologically active portions of immunoglobulin molecules include F(ab) and $F(ab')_2$ fragments which can be generated by treating the antibody with an enzyme such as pepsin. The invention provides polyclonal and monoclonal antibodies that bind to marker proteins. The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, includes a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope. A monoclonal antibody composition thus typically displays a single binding affinity for a particular marker protein with which it immunoreacts.

[0137] Polyclonal anti-marker protein antibodies can be prepared as described above by immunizing a suitable subject with a marker protein of the invention. The anti-marker protein antibody titer in the immunized subject can be monitored over time by standard techniques, such as with an enzyme linked immunosorbent assay (ELISA) using immobilized marker protein. If desired, the antibody molecules directed against marker proteins can be isolated from the mammal (*e.g.*, from the blood) and further purified by well known techniques, such as protein A chromatography, to obtain the IgG fraction. At an appropriate time after immunization, *e.g.*, when the anti-marker protein antibody titers are highest, antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma technique originally described by Kohler and Milstein (1975) *Nature* 256:495-497) (see also, Brown *et al.* (1981) *J. Immunol.* 127:539-46; Brown *et al.* (1980) *J. Biol. Chem.* 255:4980-83; Yeh *et al.*

(1976) *Proc. Natl. Acad, Sci. USA* 76:2927-3I; and Yeh *et al.* (1982) *Int. J. Cancer* 29:269-75), the more recent human B cell hybridoma technique (Kozbor *et al.* (1983) *Immunol Today* 4:72), the EBV-hybridoma technique (Cole *et al.* (1985), *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96) or trioma techniques. The technology for producing monoclonal antibody hybridomas is well known (see generally R. H. Kenneth, in *Monoclonal Antibodies: A New Dimension In Biological Analyses*, Plenum Publishing Corp., New York, New York (1980); E. A. Lerner (1981) *Yale J. Biol. Med.*, 54:387-402; M.L. Gefter *et al.* (1977) *Somatic Cell Genet.* 3:231-36). Briefly, an immortal cell line (typically a myeloma) is fused to lymphocytes (typically splenocytes) from a mammal immunized with a marker protein immunogen as described above, and the culture supernatants of the resulting hybridoma cells are screened to identify a hybridoma producing a monoclonal antibody that binds to a marker protein of the invention.

[0138] Any of the many well known protocols used for fusing lymphocytes and immortalized cell lines can be applied for the purpose of generating an anti-marker protein monoclonal antibody (see, *e.g.*, G. Galfre *et al.* (1977) *Nature* 266:SSOS2; Gefter *et al. Somatic Cell Genet.*, cited *supra*; Letter, *Yale J. Biol. Med.*, cited *supra*; Kenneth, *Monoclonal Antibodies*, cited *supra*). Moreover, the ordinarily skilled worker will appreciate that there are many variations of such methods which also would be useful. Typically, the immortal cell line (*e.g.*, a myeloma cell line) is derived from the same mammalian species as the lymphocytes. For example, murine hybridomas can be made by fusing

lymphocytes from a mouse immunized with an immunogenic preparation of the present invention with an immortalized mouse cell line. Preferred immortal cell lines are mouse myeloma cell lines that are sensitive to culture medium containing hypoxanthine, axninopterin and thymidine ("HAT medium"). Any of a number of myeloma cell lines can be used as a fusion partner according to standard techniques, *e.g.*, the P3-NS1/1-Ag4-1, P3-x63-Ag8.653 or Sp210-Ag14 myeloma lines. These myeloma lines are available from ATCC. Typically, HAT-sensitive mouse myeloma cells are fused to mouse splenocytes using polyethylene glycol ("PEG"). Hybridoma cells resulting from the fusion are then selected using HAT medium, which kills unfused and unproductively fused myeloma cells (unfused splenocytes die after several days because they are not transformed). Hybridoma cells producing a monoclonal antibody of the invention are detected by screening the hybridoma culture supernatants for antibodies that bind to a marker protein, *e.g.*, using a standard ELISA assay.

[0139] Alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal anti-marker protein antibody can be identified and isolated by screening a recombinant combinatorial immunoglobulin library (*e.g.*, an antibody phase display library) with marker protein to thereby isolate immunoglobulin library members that bind to a marker protein. Kits for generating and screening phage display libraries are commercially available (*e.g.*, the Pharmacia *Recombinant Phage Antibody System*, Catalog No. 27-9400-01; and the Stratagene *SurfZAP™ Phage Display Kit*, Catalog No. 240612). Additionally, examples of methods and

reagents particularly amenable for use in generating and screening antibody display library can be found in, for example, Ladner *et al.* U.S. Patent No. 5,223,409; Kang *et al.* PCT International Publication No. WO 92/18619; Dower *et al.* PCT International Publication No. WO 91/17271; Winter *et al.* PCT International Publication WO 92/20791; Markland *et al.* PCT International Publication No. WO 92115679; Breitling *et al.* PCT International Publication WO 93/01288; McCafferty *et al.* PCT International Publication No. WO 92/01047; Garrard *et al.* PCT International Publication No. WO 92/09690; Ladner *et al.* PCT International Publication No. WO 90/02809; Fuchs *et al.* (1991) *Bio/Technology* 9:1370-1372; Hay *et al.* (1992) *Hum. Antibod Hybridomas* 3:81-85; Huse *et al.* (1989) *Science* 246:1275-1281; Griffiths *et al.* (1993) *EMBO J* 12:725-734; Hawkins *et al.* (1992) *J. Mol. Biol.* 226:889-896; Clarkson *et al.* (1991) *Nature* 352:624-628; Gram *et al..*(1992) *Proc. Natl. Acad Sci. USA* 89:3576-3580; Garrad *et al.* (1991) *Bio/Technology* 9:1373-1377; Hoogenboom *et al.* (1991) *Nuc. Acid Res.* 19:4133-4137; Barbas *et al.* (1991) *Proc. Natl. Acad. Sci. USA* 88:7978-7982; and McCafferty *et al. Nature* (1990) 348:552-554.

[0140] Additionally, recombinant anti-marker protein antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in Robinson *et al.* International Application No. PCT/US86/02269;

Akira, *et al.* European Patent Application 184,187; Taniguchi, M., European Patent Application 171,496; Morrison *et al.* European Patent Application 173,494; Neuberger *et al.* PCT International Publication No. WO 86/01533; Cabilly *et al.* U.S. Patent No. 4,816,567; Cabilly *et al.* European Patent Application 125,023; Better *et al.* (1988) *Science* 240:1041-1043; Liu *et al.* (1987) *Proc. Natl. Acad Sci. USA* 84:3439-3443; Liu *et al.* (1987) *J. Immunol.* 139:3521 3526; Sun *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:214-218; Nishimura *et al.* (1987) *Canc. Res.* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-449; and Shaw *et al.* (1988) *J. Natl. Cancer Inst.* 80:1553-1559); Morrison, S.L. (1985) *Science* 229:1202-1207; Oi *et al.* (1986) *BioTechniques* 4:214; Winter U.S. Patent 5,225,539; Jones et al. (1986) *Nature* 321:552-525; Verhoeyan *et al.* (1988) *Science* 239:1534; and Beidler *et al.* (1988) *J. Immunol.* 141:4053-4060.

[0141] Humanized antibodies are particularly desirable for therapeutic treatment of human subjects. Humanized forms of non-human (e.g. murine) antibodies are chimeric molecules of immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', $F(ab')_2$ or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues forming a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by

corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immmunoglobulin and all or substantially all of the constant regions being those of a human immunoglobulin consensus sequence. The humanized antibody will preferably also comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al. *Nature* 321: 522-525 (1986); Riechmann et al, *Nature* 323: 323-329 (1988); and Presta *Curr.Op.Struct.Biol.* 2: 594-596 (1992).

[0142] Such humanized antibodies can be produced using transgenic mice which are incapable of expressing endogenous immunoglobulin heavy and light chains genes, but which can express human heavy and light chain genes. The transgenic mice are immunized in the normal fashion with a selected antigen, *e.g.*, all or a portion of a polypeptide corresponding to a marker of the invention. Monoclonal antibodies directed against the antigen can be obtained using conventional hybridoma technology. The human immunoglobulin transgenes harbored by the transgenic mice rearrange during B cell differentiation, and subsequently undergo class switching and somatic mutation. Thus, using such a technique, it is possible to produce therapeutically useful IgG, IgA and IgE antibodies. For an overview of this technology for producing humanized antibodies, see Lonberg and Huszar (1995) *Int. Rev. Immunol.* 13:65-93). For a detailed discussion of this technology for

producing humanized antibodies and humanized monoclonal antibodies and protocols for producing such antibodies, see, *e.g.*, U.S. Patent 5,625,126; U.S. Patent 5,633,425; U.S. Patent 5,569,825; U.S. Patent 5,661,016; and U.S. Patent 5,545,806. In addition, companies such as Abgenix, Inc. (Freemont, CA), can be engaged to provide humanized antibodies directed against a selected antigen using technology similar to that described above.

[0143] Humanized antibodies which recognize a selected epitope can be generated using a technique referred to as "guided selection." In this approach a selected non-human monoclonal antibody, *e.g.*, a murine antibody, is used to guide the selection of a humanized antibody recognizing the same epitope (Jespers *et al.*, 1994, *Bio/technology* 12:899-903).

[0144] An anti-marker protein antibody (*e.g.*, monoclonal antibody) can be used to isolate a marker protein of the invention by standard techniques, such as affinity chromatography or immunoprecipitation. An anti-marker protein antibody can facilitate the purification of natural marker proteins from cells and of recombinantly produced marker proteins expressed in host cells. Moreover, an anti-marker protein antibody can be used to detect marker protein (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the marker protein. Anti-marker protein antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling

(*i.e.*, physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatasc, galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I, ^{131}I, ^{35}S or ^{3}H.

III. Recombinant Expression Vectors and Host Cells

[0145] Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding a marker protein of the invention (or a portion thereof). As used herein, the term "vector" includes a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which includes a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and

episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

[0146] The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequences) in a manner which allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Such

regulatory sequences are described, for example, in Goeddel; *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in many types of host cells and those which direct expression of the nucleotide sequence only in certain host cells (*e.g.*, tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, and the like. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (*e.g.*, marker proteins, mutant forms of marker proteins, fusion proteins, and the like).

[0147] The recombinant expression vectors of the invention can be designed for expression of marker proteins in prokaryotic or eukaryotic cells. For example, marker proteins can be expressed in bacterial cells such as *E. coli*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

[0148] Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing

constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith, D,B. and Johnson, K.S. (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRITS (Pharmacia, Piscataway, NJ) which fuse glutathione S transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

[0149] Purified fusion proteins can be utilized in marker activity assays, (*e.g.*, direct assays or competitive assays described in detail below), or to generate antibodies specific for marker proteins, for example.

[0150] Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Hmann *et al.*, (1988) *Gene* 69:301-315) and pET 11d (Studier *et al.*,

Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, California (1990) 60-89). Target gene expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gn10-lac fusion promoter mediated by a coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter.

[0151] One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, S., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128). Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wade *et al.*, (1992) *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

[0152] In another embodiment, the marker protein expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1 (Baldari, *et al.*, (1987) *Embo J.* 6:229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz *et al.*, 21987) *Gene*

54:113-123), pYES2 (InVitrogen Corporation, San Diego, CA), and picZ (Invitrogen Corp, San Diego, CA).

[0153] Alternatively, marker proteins of the invention can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.*, Sf 9 cells) include the pAc series (Smith *et al.* (1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

[0154] In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, B. (1987) *Nature* 329:840) and pMT2PC (Kaufman *et al.* (1987) *EMBO J.* 6:187-I95). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed.. Cold Spring Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 60-89). Target gene expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gn10-lac fusion promoter mediated by a coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a

resident prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter.

[0155] One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, S., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128). Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, (1992) *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

[0156] In another embodiment, the marker protein expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSecl (Baldari, *et al.*, (1987) *Embo J.* 6:229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz et al., (1987) *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, CA), and picZ (InVitrogen Corp, San Diego, CA).

[0157] Alternatively, marker proteins of the invention can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.*, Sf9 cells) include the pAc series (Smith *et al.* (1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

[0158] In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, B. (1987) *Nature* 329:840) and pMT2PC (Kaufman *et al.* (1987) *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed, Cold Spring Harbor Laboratory*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

[0159] In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.* (1987) *Genes Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv. Immunol.* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J.* 8:729-733) and immunoglobulins (Banerji *et al.* (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters (*e.g.*, the neurofilament promoter, Byrne and R.aaddle (1989) *Proc.*

Nall. Acad Sci. USA 86:5473-5477), pancreas-specific promoters (Edlund *et al.* (1985) *Science* 230:912-916), and mammary gland-specific promoters (*e.g.*, milk whey promoter, U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, for example the marine hox promoters (Kessel and Grass (I990) *Science* 249:374-379) and the α-fetoprotein promoter (Campes and Tilghman (1989) *Genes Dev.* 3:537-546).

[0160] The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner which allows for expression (by transcription of the DNA molecule) of an RNA molecule which is antisense to mRNA corresponding to a gene of the invention (*e.g.*, a gene set forth in Tables 1-5 or 7-10). Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub, H. *et al.*, Antisense RNA as a

molecular tool for genetic analysis, *Reviews - Trends in Genetics*, Vol. 1(1)1986.

[0161] Another aspect of the invention pertains to host cells into which a nucleic acid molecule of the invention is introduced, *e.g.*, a gene set forth in Tables 1-5 or 7-10 within a recombinant expression vector or a nucleic acid molecule of the invention containing sequences which allow it to homologously recombine into a specific site of the host cell's genome. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

[0162] A host cell can be any prokaryotic or eukaryotic cell. For example, a marker protein of the invention can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

[0163] Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for

introducing foreign nucleic acid (*e.g.*, DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electmporation. Suitable methods for transforming or transferring host cells can be found in Sambrook, *et al.* (*Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY,1989), and other laboratory manuals.

[0164] For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable flag (*e.g.*, resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable flags include those which confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable flag can be introduced into a host cell on the same vector as that encoding a marker protein or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (*e.g.*, cells that have incorporated the selectable flag gene will survive, while the other cells die).

[0165] A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (*i.e.*, express) a marker protein. Accordingly, the invention further provides methods for producing a marker protein using the host cells of the

invention. In one embodiment, the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding a marker protein has been introduced) in a suitable medium such that a marker protein of the invention is produced. In another embodiment, the method further comprises isolating a marker protein from the medium or the host cell.

[0166] The host cells of the invention can also be used to produce non-human transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which marker-protein-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous sequences encoding a marker protein of the invention have been introduced into their genome or homologous recombinant animals in which endogenous sequences encoding the marker proteins of the invention have been altered. Such animals are useful for studying the function and/or activity of a marker protein and for identifying and/or evaluating modulators of marker protein activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, and the tike. A transgene is exogenous DNA which is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic

animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous gene of the invention (*e.g.*, a gene set forth in Tables 1-5 or 7-10) has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, *e.g.*, an embryonic cell of the animal, prior to development of the animal.

[0167] A transgenic animal of the invention can be created by introducing a marker-encoding nucleic acid into the mate pronuclei of a fertilized oocyte, *e.g.*, by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to a transgene to direct expression of a marker protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009, both by Leder *et al.*, U.S. Patent No. 4,873,191 by Wagner *et al.* and in Hogan, B., *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of a transgene of the invention in its genome and/or expression of mRNA corresponding to a gene of the invention in tissues or cells of the animals. A transgenic founder animal can then be used to breed

additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding a marker protein can further be bred to other transgenic animals carrying other transgenes.

[0168] To create a homologous recombinant animal, a vector is prepared which contains at least a portion of a gene of the invention into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the gene. The gene can be a human gene, but more preferably, is a non-human homologue of a human gene of the invention (*e.g.*, a gene set forth in Tables 1-5 or 7-10). For example, a mouse gene can be used to construct a homologous recombination nucleic acid molecule, *e.g.*, a vector, suitable far altering an endogenous gene of the invention in the mouse genome. In a preferred embodiment, the homologous recombination nucleic acid molecule is designed such that, upon homologous recombination, the endogenous gene of the invention is functionally disrupted (*i.e.*, no longer encodes a functional protein; also referred to as a "knock out" vector). Alternatively, the homologous recombination nucleic acid molecule can be designed such that, upon homologous recombination, the endogenous gene is mutated or otherwise altered but still encodes functional protein (*e.g.*, the upstream regulatory region can be altered to thereby alter the expression of the endogenous marker protein). In the homologous recombination nucleic acid molecule, the altered portion of the gene of the invention is flanked at its 5' and 3' ends by additional nucleic acid sequence of the gene of the invention to allow for homologous recombination to occur between the exogenous gene carried by the homologous recombination nucleic acid

molecule and an endogenous gene in a cell, *e.g.*, an embryonic stem cell. The additional flanking nucleic acid sequence is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the homologous recombination nucleic acid molecule (see, *e.g.*, Thomas, K.R. and Capecchi, M.R. (1987) *Cell* 51:503 for a description of homologous recombination vectors). The homologous recombination nucleic acid molecule is introduced into a cell, *e.g.*, an embryonic stem cell line (*e.g.*, by electroporation) and cells in which the introduced gene has homologously recombined with the endogenous gene are selected (see *e.g.*, Li, E. *et al.* (1992) *Cell* 69:915). The selected cells can then injected into a blastocyst of an animal (*e.g.*, a mouse) to form aggregation chimeras (see *e.g.* Bradley, S A. in *Teratocareirtomas and Embryonic Stem Cells: A Practical Approach*, E.J. Robertson, ed. (IRL, Oxford, 1987) pp. 113-152). A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination nucleic acid molecules, *e.g.*, vectors, or homologous recombinant animals are described further in Bradley, A. (1991) *Current Opinion in Biotechnology* 2:823-829 and in PCT International Publication Nos.: WO 90/11354 by Le Mouellec *et al.*; WO 91/01140 by Smithies *et al.*; WO 92/0968 by Zijlstra *et al.*; and WO 93/04169 by Berns *et al.*

[0169] In another embodiment, transgenic non-human animals can be produced which contain selected systems which allow for regulated expression of the transgene. One example of such a system is the *cre/loxP* recombinase system of bacteriophage P1. For a description of the *cre/loxP* recombinase system, see, *e.g.*, Laksa *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman *et al.* (1991) *Science* 251:1351-1355. If a *cre/loxP* recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the *Cre* recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, *e.g.*, by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

[0170] Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut, I. *et al.* (1997) *Nature* 385:810-813 and PCT International Publication Nos. WO 97/07668 and WO 97/07669. In brief, a cell, *e.g.*, a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G_o phase. The quiescent cell can then be fused, *e.g.*, through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyte and then transferred to pseudopregnant

female foster animal. The offspring borne of this female foster animal will be a clone of the animal from which the cell, *e.g.*, the somatic cell, is isolated.

IV. Pharmaceutical Compositions

[0171] The nucleic acid molecules of the invention (*e.g.*, the genes set forth in Tables 1-5 or 7-10), fragments of marker proteins, and anti-marker protein antibodies of the invention can be incorporated into pharmaceutical compositions suitable for administration. Such compositions (also referred to herein as "bioactive agents or compounds") typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier.

[0172] As used herein the language "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. The use of such media and agents for pharmaceutically active substances is well-known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary bioactive agents can also be incorporated into the compositions.

[0173] The invention includes methods for preparing pharmaceutical compositions for modulating the expression or activity of a polypeptide or nucleic acid corresponding to a marker of the invention. Such methods comprise formulating a pharmaceutically

acceptable carrier with an agent which modulates expression or activity of a polypeptide or nucleic acid corresponding to a marker of the invention. Such compositions can further include additional active agents. Thus, the invention further includes methods for preparing a pharmaceutical composition by formulating a pharmaceutically acceptable carrier with an agent which modulates expression or activity of a polypeptide or nucleic acid corresponding to a marker of the invention and one or more additional bioactive agents.

[0174] The invention also provides methods (also referred to herein as "screening assays") for identifying modulators, i.e., candidate or test compounds or agents comprising therapeutic moieties (*e.g.*, peptides, peptidomimetics, peptoids, small molecules or other drugs) which (a) bind to the marker, or (b) have a modulatory (*e.g.*, stimulatory or inhibitory) effect on the activity of the marker or, more specifically, (c) have a modulatory effect on the interactions of the marker with one or more of its natural substrates (*e.g.*, peptide, protein, hormone, co-factor, or nucleic acid), or (d) have a modulatory effect on the expression of the marker. Such assays typically comprise a reaction between the marker and one or more assay components. The other components may be either the test compound itself, or a combination of test compound and a natural binding partner of the marker.

[0175] The test compounds of the present invention may be bioactive agents, i.e. protein, oligopeptide, molecule, polysaccharide, polynucleotides. In a preferred embodiment the bioactive agents are proteins,

in particular naturally occurring proteins or fragments thereof.

[0176] The test compounds of the present invention may be obtained from any available source, including systematic libraries of natural and/or synthetic compounds. Test compounds may also be obtained by any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; peptoid libraries (libraries of molecules having the functionalities of peptides, but with a novel, non-peptide backbone which are resistant to enzymatic degradation but which nevertheless remain bioactive; see, *e.g.*, Zuckermann *et al.*, 1994, *J. Med. Chem.* 37:2678-85); spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the 'one-bead one-compound' library method; and synthetic library methods using affinity chromatography selection. The biological library and peptoid library approaches are limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, 1997, *Anticancer Drug Des.* 12:145).

[0177] A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, *e.g.*, intravenous, intradermal, subcutaneous, oral (*e.g.*, inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile

diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine; propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfate; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

[0178] Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The earner can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyetheylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the requited particle size in the case of dispersion and by the use

of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

[0179] Sterile injectable solutions can be prepared by incorporating the active compound (*e.g.*, a fragment of a marker protein or an anti-marker protein antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enmnerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active, ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

[0180] Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used

in the form of tablets, troches, or capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Stertes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

[0181] For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer.

[0182] Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the

bioactive compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

[0183] The compounds can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

[0184] In one embodiment, the therapeutic moieties, which may contain a bioactive compound, are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

[0185] It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein includes physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity

of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on-the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

[0186] Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD50 (the dose lethal to 50% of the population) and the ED50 (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD50/ED50. Compounds which exhibit large therapeutic indices are preferred. While compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

[0187] The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically

effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC50 (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

[0188] The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (see U.S. Patent 5,328,470) or by stereotactic injection (see *e.g.*, Chen *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, *e.g.*, retroviral vectors, the pharmaceutical preparation can include one or more cells which produce the gene delivery system.

[0189] The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

V. Computer Readable Means and Arrays

[0190] Computer readable media comprising a marker(s) of the present invention is also provided. As used herein, "computer readable media" includes a medium that can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. The skilled artisan will readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a marker of the present invention.

[0191] As used herein, "recorded" includes a process for storing information on computer readable medium. Those skilled in the art can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising the markers of the present invention.

[0192] A variety of data processor programs and formats can be used to store the marker information of the present invention on computer readable medium. For example, the nucleic acid sequence corresponding to the markers can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. Any number of dataprocessor structuring formats (*e.g.*,

text file or database) may be adapted in order to obtain computer readable medium having recorded thereon the markers of the present invention.

[0193] By providing the markers of the invention in computer readable form, one can routinely access the marker sequence information for a variety of purposes. For example, one skilled in the art can use the nucleotide or amino acid sequences of the invention in computer readable form to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of the sequences of the invention which match a particular target sequence or target motif.

[0194] The invention also includes an array comprising a marker(s) of the present invention, i.e. a biochip. The array can be used to assay expression of one or more genes in the array. In one embodiment, the array can be used to assay gene expression in a tissue to ascertain tissue specificity of genes in the array. In this manner, up to about 8600 genes can be simultaneously assayed for expression. This allows an expression profile to be developed showing a battery of genes specifically expressed in one or more tissues at a given point in time.

[0195] In addition to such qualitative determination, the invention allows the quantitation of gene expression in the biochip. Thus, not only tissue specificity, but also the level of expression of a battery of genes in the tissue is ascertainable. Thus, genes can be grouped on the basis of their tissue expression per se and level of expression in that

tissue. As used herein, a "normal level of expression" refers to the level of expression of a gene provided in a control sample, typically the control is from non-involved cells or tissues, or from a non-diseased subject. The determination of normal levels of expression is useful, for example, in ascertaining the relationship of gene expression between or among tissues. Thus, one tissue can be perturbed and the effect on gene expression in a second tissue can be determined. In this context, the effect of one cell type on another cell type in response to a biological stimulus can be determined. Such a determination is useful, for example, to know the effect of cell-cell interaction at the level of gene expression. If an agent is administered therapeutically to treat one cell type but has an undesirable effect on another cell type, the invention provides an assay to determine the molecular basis of the undesirable effect and thus provides the opportunity to co-administer a counteracting agent or otherwise treat the undesired effect. Similarly, even within a single cell type, undesirable biological effects can be determined at the molecular level. Thus, the effects of an agent on expression of other than the target gene can be ascertained and counteracted.

[0196] In another embodiment, the arrays can be used to monitor the time course of expression of one or more genes in the array. This can occur in various biological contexts, as disclosed herein, for example development and differentiation, disease progression, *in vitro* processes, such a cellular transformation and senescence, autonomic neural and neurological processes, such as, for example, pain and appetite, and cognitive functions, such as learning or memory.

[0197] The array is also useful for ascertaining the effect of the expression of a gene on the expression of other genes in the same cell or in different cells. This provides, for example, for a selection of alternate molecular targets for therapeutic intervention if the ultimate or downstream target cannot be regulated.

[0198] The array is also useful for ascertaining differential expression patterns of one or more genes in non-involved or diseased cells. This provides a battery of genes that could serve as a molecular target for diagnosis or therapeutic intervention. In particular, biochips can be made comprising arrays not only of the differentially expressed markers listed in Tables 1-5 or 7-10, but of markers specific to subjects at a certain stage of the disease (i.e. secondary progressive, primary progressive, relapsing-remitting), or from higher-risk geographical regions (i.e. northern Europe or Canada).

VI. Predictive Medicine

[0199] The present invention pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, pharmacogenetics and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining marker protein and/or nucleic acid expression as welt as marker protein activity, in the context of a biological sample (*e.g.*, blood, serum,

cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with increased or decreased marker protein expression or activity. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with marker protein, nucleic acid expression or activity. For example, the number of copies of a marker gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purposes to thereby phophylactically treat an individual prior to the onset of a disorder (*e.g.*, multiple sclerosis)characterized by or associated with marker protein, nucleic acid expression or activity.

[0200] Another aspect of the invention pertains to monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of marker in clinical trials.

[0201] These and other agents are described in further detail in the following sections.

1. Diagnostic Assays

[0202] An exemplary method for detecting the presence or absence of marker protein or nucleic acid of the invention in a biological sample involves obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting the protein or nucleic acid (*e.g.*, mRNA, genomic DNA) that encodes the marker protein such that the presence of the marker protein or nucleic acid

is detected in the biological sample. A preferred agent for detecting mRNA or genomic DNA corresponding to a marker gene or protein of the invention is a labeled nucleic acid probe capable of hybridizing to a mRNA or genomic DNA of the invention. Suitable probes for use in the diagnostic assays of the invention are described herein.

[0203] A preferred agent for detecting marker protein is an antibody capable of binding to marker protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (*e.g.*, Fab or $F(ab')_2$) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (*i.e.*, physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to defeat marker mRNA, protein, or genomic DNA in a biological sample *in vitro* as well as *in vivo*. For example, in vitro techniques for detection of marker mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of marker protein include enzyme linked immunosorbent assays (ELISAs),

Western blots, immunoprecipitations and immunofluoresCence. *In vitro* techniques for detection of marker genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of marker protein include introducing into a subject a labeled anti-marker antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

[0204] In one embodiment, the biological sample contains protein molecules from the test subject. Alternatively, the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject. A preferred biological sample is a serum sample isolated by conventional means from a subject.

[0205] In another embodiment, the methods further involve obtaining a control biological sample (*e.g.*, noninvolved tissue or from a non-diseased subject) from a control subject, contacting the control sample with a compound or agent capable of detecting marker protein, mRNA, or genomic DNA, such that the presence of marker protein, mRNA or genomic DNA is detected in the biological sample, and comparing the presence of marker protein, mRNA or genomic DNA in the control sample with the presence of marker protein, mRNA or genomic DNA in the test sample.

[0206] The invention also encompasses kits for detecting the presence of marker in a biological sample. For example, the kit can comprise a labeled compound or agent capable of detecting marker protein

or mRNA in a biological sample; means for determining the amount of marker in the sample; and means for comparing the amount of marker in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect marker protein or nucleic acid.

2. Prognostic Assays

[0207] The diagnostic methods, described herein can furthermore be utilized to identify subjects having or at risk of developing a disease or disorder associated with aberrant marker expression or activity. As used herein, the term "aberrant" includes a marker expression or activity which deviates from the wild type marker expression or activity. Aberrant expression or activity includes increases or decreased expression or activity, as well as expression or activity which does not follow the wild type developmental pattern of expression or the subcellular pattern of expression. For example, aberrant marker expression or activity is intended to include the cases in which a mutation in the marker gene causes the marker gene to be under-expressed or over-expressed and situations in which such mutations result in a non-functional marker protein or a protein which does not function in a wild-type fashion, *e.g.*, a protein which does not interact with a marker ligand or one which interacts with a non marker protein ligand.

[0208] The assays described herein, such as the preceding diagnostic assays or the following assays, can be utilized to identify a subject having or at risk of developing a disorder associated with a

misregulation in marker protein activity or nucleic acid expression, such as multiple sclerosis. Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing a, disorder associated with a misregulation in marker protein activity or nucleic acid expression, such as multiple sclerosis. Thus, the present invention provides a method for identifying a disease or disorder associated with aberrant marker expression or activity in which a test sample is obtained from a subject and marker protein or nucleic acid (*e.g.*, mRNA or genomie DNA) is detected, wherein the presence of marker protein or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant marker expression or activity. As used herein, a "test sample" includes a biological sample obtained from a subject of interest. For example, a test sample can be a biological fluid (*e.g.*, blood PBMCs), cell sample, or tissue (*e.g.*, brain).

[0209] Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with increased or degreased marker expression or activity. For example, such methods can be used to determine whether a subject can be effectively treated with an agent for a disorder such as multiple sclerosis. Thus, the present invention provides methods for determining whether a subject can be effectively treated with an agent for a disorder associated with increased or decreased marker expression or activity in which a test sample is obtained and marker protein or nucleic acid expression

or activity is detected (*e.g.*, wherein the abundance of marker protein or nucleic acid expression or activity is diagnostic for a subject that can be administered the agent to treat a disorder associated with increased or decreased marker expression or activity).

[0210] The methods of the invention can also be used to detect genetic alterations in a marker gene, thereby determining if a subject with the altered gene is at risk for a disorder characterized by misregulation in marker protein activity or nucleic acid expression, such as multiple sclerosis. In preferred embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic alteration characterized by at least one of an alteration affecting the integrity of a gene encoding a marker-protein, or the mis-expression of the marker gene. For example, such genetic alterations can be detected by ascertaining the existence of at least one of 1) a deletion of one or more nucleotides from a marker gene; 2) an addition of one or more nucleotides to a marker gene; 3) a substitution of one or more nucleotides of a marker gene, 4) a chromosomal rearrangement of a marker gene; 5) an alteration in the level of a messenger RNA transcript of a marker gene, 6) aberrant modification of a ma=ker gene, such as of the methylation pattern of the genomic DNA, 7) the presence of a non-wild type splicing pattern of a messenger RNA transcript of a marker gene, 8) a non-wild type level of a marker-protein, 9) allelic loss of a marker gene, and 10) inappropriate post-translational modification of a marker-protein. As described herein, there are a large number of assays known in the art which can be used for detecting alterations in a marker gene. A preferred biological sample is a tissue (*e.g.*,

brain) or blood sample isolated by conventional means from a subject.

[0211] In certain embodiments, detection of the alteration involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, *e.g.*, U.S. Patent Nos. 4,683,(95 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, *e.g.*, Landegran *et al.* (1988) *Science* 241:1077-1080; and Nakazawa *et al.* (1994) *Proc. Mail. Acad. Sci. USA* 91:360-364), the latter of which can be particularly useful for detecting point mutations in the marker-gene (see Abravaya *et al.* (1995) *Nucleic Acids Res.* 23:675-682). This method can include the steps of collecting a sample of cells from a subject, isolating nucleic acid (*e.g.*, genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to a marker gene under conditions such that hybridization and amplification of the marker-gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

[0212] Alternative amplification methods include: self sustained sequence replication (Guatelli, JC. *et al.*, (1990) *Proc. Natl. Acad. Sci. USA* 87:1874-1878), transcriptional amplification system (Kwoh, D.Y. *et al.*, (1989) *Proc. Natl. Acad. Sci. USA* 86:1173-1177),

Q-Beta Replicase (Lizardi, P.M. *et al.* (1988) *Bio-Technology* 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

[0213] In an alternative embodiment, mutations in a marker gene from a sample cell can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis and compared. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (see, for example, U.S. Patent No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

[0214] In other embodiments, genetic mutations in a marker gene or a gene encoding a marker protein of the invention can be identified by hybridizing a sample and control nucleic acids, *e.g.*, DNA or RNA, to high density arrays containing hundreds or thousands of oligonucleotides probes (Cronin, M.T. *et al.* (1996) *Human Mutation* 7: 244-255; Kozal, M.J. *et al.* (1996) *Nature Medicine* 2: 753-759). For example, genetic mutations in marker can be identified in two dimensional arrays containing light generated DNA probes as described in Cronin, M.T. *et al. supra*.

Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This step is followed by a second hybridization array that allows the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

[0215] In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the marker gene and detect mutations by comparing the sequence of the sample marker with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxam and Gilbert ((1977) *Proc. Natl. Acad. Scl. USA* 74:560) or Sanger ((1977) *Proc. Natl. Acad. Sci. USA* 74:5463). It is also contemplated that any of a variety of automated sequencing procedures can be utilized when performing the diagnostic assays ((1995) *Biotechniques* 19:448), including sequencing by mass spectrometry (see, *e.g.*, PCT International Publication No. WO 94116101; Cohen *et al.* (1996) *Adv. Chromatogr.* 36:127-162; and Griffin *et al.* (1993) *Appl. Biochem. Biotechnol.* 38:147-159).

[0216] Other methods for detecting mutations in the marker gene or gene encoding a marker protein of the invention include methods in which protection from

cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA heteroduplexes (Myers *et al.* (1985) *Science* 230:1242). In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes by hybridizing (labeled) RNA or DNA containing the wild-type marker sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent which cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digest the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. See, for example, Cotton *et al.* (1988) *Proc. Natl Acad Sci USA* 85:4397; Saleeba *et al.* (1992) *Methods Enzymol.* 517:286-295. In a preferred embodiment, the control DNA or RNA can be labeled for detection.

[0217] In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in marker cDNAs obtained from samples of cells. For example, the mutY enzyme of *E. coli* cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches (Hsu *et al.* (1994)

Carcinogenesis 15:1657-1652). According to an exemplary embodiment, a probe based on a marker sequence, *e.g.*, a wild-type marker sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. See, for example, U.S. Patent No. 5,459,039.

[0218] In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in marker genes or genes encoding a marker protein of the invention. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids (Orita *et al.* (1989) *Proc Natl. Acad. Sci USA*: 86:2766, see also Cotton (1993) *Mutat. Res.* 285:125-144; and Hayashi (1992) *Genet. Anal. Tech Appl.* 9:73-79). Single-stranded DNA fragments of sample and control marker nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In a preferred embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in elecrtophoretic mobility (Keen *et al.* (1991) *Trends Genet* 7:5).

[0219] In yet another embodiment the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE) (Myers *et al.* (1985) *Nature* 313:495). When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 by of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA (Rosenbaum and Reissner (1987) *Biophys Chem* 265:12753).

[0220] Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions which permit hybridization only if a perfect match is found (Saiki *et al.* (1986) *Nature* 324:163); Saiki *et al.* (1989) *Proc. Natl. Acad. Sci USA* 86:6230). Such allele specific oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

[0221] Alternatively, allele specific amplification technology which depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific

amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs *et al.* (1989) *Nucleic Acids Res.* 17:2437-2448) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (Prossner (1993) *Tibtech* 11:238). In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection (Gasparini *et al.* (1992) *Mol. Cell Probes* 6:1). It is anticipated that in certain embodiments amplification may also be performed using Taq ligase for amplification (Barany (1991) *Proc. Natl. Acad. Sci USA* 88:189). In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

[0222] The methods described herein may be performed, for example, by utilizing prepackaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which may be conveniently used, *e.g.*, in clinical settings to diagnose subjects exhibiting symptoms or family history of a disease or illness involving a marker gene.

[0223] Furthermore, any cell type or tissue in which marker is expressed may be utilized in the prognostic assays described herein.

3. Monitoring of Effects During Clinical Trials

[0224] Monitoring the influence of agents (*e.g.*, drugs) on the expression or activity of a marker protein (*e.g.*, the modulation of multiple sclerosis) can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent determined by a screening assay as described herein to increase marker gene expression, protein levels, or upregulate marker activity, can be monitored in clinical trials of subjects exhibiting decreased marker gene expression, protein levels, or downruegulated marker activity. Alternatively, the effectiveness of an agent determined by a screening assay to decrease marker gene expression, protein levels, or downregulate marker activity, can be monitored in clinical trials of subjects exhibiting increased marker gene expression, protein levels, or upregulated marker activity. In such clinical trials, the expression or activity of a marker gene, and preferably, other genes that have been implicated in, for example, a marker-associated disorder (*e.g.*, multiple sclerosis) can be used as a "read out" or markers of the phenotype of a particular cell.

[0225] For example, and not by way of limitation, genes, including marker genes and genes encoding a marker protein of the invention, that are modulated in cells by treatment with an agent (*e.g.*, compound, drug or small molecule) which modulates marker activity (*e.g.*, identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents on marker-associated disorders (*e.g.*,

multiple sclerosis), for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of marker and other genes implicated in the marker-associated disorder, respectively. The levels of gene expression (*e.g.*, a gene expression pattern) can be quantified by northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods as described herein, or by measuring the levels of activity of marker or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state may be determined before, and at various points during treatment of the individual with the agent.

[0226] In a preferred embodiment, the present invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate identified by the screening assays described herein) including the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of a marker protein, mRNA, or genomic DNA in the pre-administration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of the marker protein, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the marker protein, mRNA, or genomic DNA in the pre-administration sample with the marker protein, mRNA, or genomic DNA in the post

administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of marker to higher levels than detected, *i.e.*, to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of marker to lower levels than detected, *i.e.* to decrease the effectiveness of the agent. According to such an embodiment, marker expression or activity may be used as an indicator of the effectiveness of an agent, even in the absence of an observable phenotypic response.

C. Methods of Treatment

[0227] The present invention provides for both prophylactic and therapeutic methods of treating a subject at risk for (or susceptible to) a disorder or having a disorder associated with aberrant marker expression or activity. With regards to both prophylactic and therapeutic methods of treatment, such treatments may be specifically tailored or modified, based on knowledge obtained from the field of pharmacogenomics. "Pharmacogenomics", as used herein, includes the application of genomics technologies such as gene sequencing, statistical genetics, and gene expression analysis to drugs in clinical development and on the market. More specifically, the term refers the study of how a subject's genes determine his or her response to a drug (*e.g.*, a subject's "drug response phenotype", or "drug response genotype".) Thus, another aspect of the invention provides methods for tailoring an individual's prophylactic or therapeutic treatment with either the marker molecules of the

present invention or marker modulators according to that individual's drug response genotype. Pharmacogenomics allows a clinician or physician to target prophylactic or therapeutic treatments to subjects who will most benefit from the treatment and to avoid treatment of subjects who will experience toxic drug-related side effects.

1. Prophylactic Methods

[0228] In one aspect, the invention provides a method for preventing in a subject, a disease or condition (*e.g.*, multiple sclerosis) associated with increased or decreased marker expression or activity, by administering to the subject a marker protein or an agent which modulates marker protein expression or at least one marker protein activity. Subjects at risk for a disease which is caused or contributed to by increased or decreased marker expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the differential marker protein expression, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of marker aberrancy (*e.g.*, increase or decrease in expression level), for example, a marker protein, marker protein agonist or marker protein antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

2. Therapeutic Methods

[0229] Another aspect of the invention pertains to methods of modulating marker protein expression or activity for therapeutic purposes. Accordingly, in an exemplary embodiment, the modulatory method of the invention involves contacting a cell with a marker protein or agent that modulates one or more of the activities of a marker protein activity associated with the cell. An agent that modulates marker protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring target molecule of a marker protein (*e.g.*, a marker protein substrate), a marker protein antibody, a marker protein agonist or antagonist, a peptidomimetic of a marker protein agonist or antagonist, or other small molecule. In one embodiment, the agent stimulates one or more marker protein activities. Examples of such stimulatory agents include active marker protein and a nucleic acid molecule encoding marker protein that has been introduced into the cell. In another embodiment, the agent inhibits one or more marker protein activities. Examples of such inhibitory agents include antisense marker protein nucleic said molecules, anti-marker protein antibodies, and marker protein inhibitors. These modulatory methods can be performed *in vitro* (*e.g.*, by culturing the cell with the agent) or, alternatively, *in vivo* (*e.g.*, by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant expression or activity of a marker protein or nucleic acid molecule. In one embodiment, the method involves administering an agent (*e.g.*, an agent identified by a screening assay described herein), or combination of

agents that modulates (*e.g.*, upregulates or downregulates) marker protein expression or activity. In another embodiment, the method involves administering a marker protein or nucleic acid molecule as therapy to compensate for reduced or aberrant marker protein expression or activity.

[0230] Stimulation of marker protein activity is desirable in situations in which marker protein is abnormally downregulated and/or in which increased marker protein activity is likely to have a beneficial effect. For example, stimulation of marker protein activity is desirable in situations in which a marker is downregulated and/or in which increased marker protein activity is likely to have a beneficial erect. Likewise, inhibition of marker protein activity is desirable in situations in which marker protein is abnormally upregulated and/or in which decreased marker protein activity is likely to have a beneficial effect.

3. Pharmacogenomics

[0231] The marker protein and nucleic acid molecules of the present invention, as well as agents, inhibitors or modulators which have a stimulatory or inhibitory effect on marker protein activity (*e.g.*, marker gene expression) as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) marker-associated disorders (*e.g.*, multiple sclerosis) associated with aberrant marker protein activity. In conjunction with such treatment, pharmacogenomics (*i.e.*, the study of the relationship between an individual's genotype and that individual's response to

a foreign compound or drug) may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, a physician or clinician may consider applying knowledge obtained in relevant pharmacogenomics studies in determining whether to administer a marker molecule or marker modulator as well as tailoring the dosage and/or therapeutic regimen of treatment with a marker molecule or marker modulator.

[0232] Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See, for example, Eichelbaum, M. *et al.* (1996) *Clin. Exp. Pharmacol. Physiol.* 23(10-11) :983-985 and Linden, M.W. *et al.* (1997) *Clin. Chem.* 43(2):254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare genetic defects or as naturally-occurring polymorphisms. For example, glucose-6-phosphate dehydrogenase deficiency (G6PD) is a common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

[0233] One pharmacogenomics approach to identifying genes that predict drug response, known as "a genome-

wide association", relies primarily on a high-resolution map of the human genome consisting of already known gene-related markers (*e.g.*, a "bi-allelic" gene marker map which consists of 60,000-100,000 polymorphic or variable sites on the human genome, each of which has two variants.) Such a high-resolution genetic map can be compared to a map of the genome of each of a statistically substantial number of subjects taking part in a Phase II/III drug trial to identify markers associated with a particular observed drug response or side effect. Alternatively, such a high resolution map can be generated from a combination of some ten-million known single nucleotide polymorphisms (SNPs) in the human genome. As used herein, a "SNP" is a common alteration that occurs in a single nucleotide base in a stretch of DNA. For example, a SNP may occur once per every 1000 bases of DNA. A SNP may be involved in a disease process, however, the vast majority may not be disease associated. Given a genetic map based on the occurrence of such SNPs, individuals can be grouped into genetic categories depending on a particular pattern of SNPs in their individual genome. In such a manner, treatment regimens can be tailored to groups of genetically similar individuals, taking into account traits that may be common among such genetically similar individuals.

[0234] Alternatively, a method termed the "candidate gene approach", can be utilized to identify genes that predict drug response. According to this method, if a gene that encodes a drugs target is known (*e.g.*, a marker protein of the present invention), all common variants of that gene can be fairly easily identified in the population and it can be determined if having

one version of the gene versus another is associated with a particular drug response.

[0235] As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (*e.g.*, N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYPZC19) has provided an explanation as to why some subjects do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. The other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

[0236] Alternatively, a method termed the "gene expression profiling", can be utilized to identify genes that predict drug response. For example, the

gene expression of an animal dosed with a drug (*e.g.*, a marker molecule or marker modulator of the present invention) can give an indication whether gene pathways related to toxicity have been turned on.

[0237] Information generated from more than one of the above pharmacogenomics approaches can be used to determine appropriate dosage and treatment regimens for prophylactic or therapeutic treatment an individual. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a marker molecule or marker modulator, such as a modulator identified by one of the exemplary screening assays described herein.

[0238] This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patents and published patent applications cited throughout this application, as well as the Figures and Tables are incorporated herein by reference.

EXAMPLES

EXAMPLE 1: IDENTIFICATION AND CHARACTERIZATION OF MARKER cDNA IN MURINE MODEL OF EXPERIMENTAL AUTOIMMUNE ENCHEPHALITIS (EAE)

A. Induction of EAE in Mice

[0239] 12 SJL mice were immunized subcutaneously with 150 μg injection of proteolipid protein (PLP), peptide residues 139-151 in complete Freud's adjuvant, a protein capable of inducing EAE in mice. After 10 days, the spleens were harvested, and splenocyte cells were isolated and reactivated *in vitro* for 4 days with 5 μg of PLP. The amplified splenocyte cells were then injected into SJL mice at 15 x 10^6 cells per mouse.

B. *In Vivo* Development of EAE and Isolation of Immune Cells and Microglial Cells

[0240] Development of EAE in mice was characterized by three stages: onset, peak and recovery. Onset of EAE in mice was characterized by a limp tail 7-8 days after injection. Peak development of EAE in mice appeared as full hind limb paralysis approximately 10-12 days after injection, while recovery appeared as hind limb weakness approximately 16-18 days after injection. At each of the above 3 stages, the following numbers of mice were sacrificed and harvested for cervical spinal cord and mid-brain: 2 mice at onset, 6 at peak and 4 at recovery. To isolate infiltrating immune cells and microglial cells, the samples were run over a percoll gradient to remove neurons, endothelial cells and astrocytes.

C. Isolation of RNA

[0241] Total RNA was isolated using the RNeasy mini kit (Quiagen, Hilden, Germany). To prepare cRNA for hybridization, 5 μg of total RNA was denatured at 70°C with T7-tagged oligo-dT primer, cooled on ice, then reverse transcribed with 200 units Superscript RT II at 50°C for 1 hour in 1x first strand buffer, 10 mM DTT and 0.5 mM of each dNTP (Gibco BRL, Gaiethersburg, MD). Second strand cDNA was synthesized by adding 40 units DNA pol I, 10 units *E. coli* DNA ligase, 2 units Rnase H, 30 μL second strand buffer, 3 μl 10 mM each dNTP, and water to 150 μL final volume and incubating at 15.8°C for 2 hours. The resulting cDNA was extracted once with phenol/chloroform/isoamylalcohol. CDNA was separated on a Phase Lock Gel tube at maximum speed for 2 min and precipitated with sodium acetate and 100 ethanol. The resulting pellet was washed with 80% ethanol, was dried and was resuspended in diethylpyrocarbonate-treated (DEPC-treated) water.

[0242] Labeled RNA was prepared from clones containing a T7 RNA polymerase promoter site by incorporating labeled ribonucleotides in an *in vitro* transcription (IVT) reaction. Half of the purified cDNA was used for *in vitro* transcription with a T7 RNA polymerase kit, following manufacturer instructions and using an overnight 37 °C incubation, thereby incorporating biotinylated CTP and UTP. Labeled RNA was purified using RNeasy columns (Quiagen). RNA was concentrated and then quantitated by spectrophotometry. Labeled RNA (13-15 μg) was fragmented in 40 mM Tris-acetate 8.0, 100 mM potassium acetate, 30 mM magnesium acetate for 35 min at 94 C in a total volume of 40 μL.

D. Array Hybridization and Detection of Fluorescence

[0243] The labeled and fragmented RNA probes were diluted in 1 x MES buffer, BIO948, Bio C, B cre, 100 μg/ml herring sperm DNA, and 50 μg/ml acetylated BSA. New probes were pre-hybridized in a microfuge tube with glass beads at 45°C overnight to remove debris. Oligonucleotide arrays composed of approximately 11,000 murine genes (Microarray, Affymetrix, Cat Nos. SubA #510243, SubB #510244) were pre-hybridized with 1 x MES hybridization buffer at 45°C for 5 min and then insoluble material was removed by centrifugation. Pre-hybridization buffer was removed from oligo array cartridges, 200 μL probe added and cartridges were hybridized for 16 hours at 45°C at 60 rpm. After hybridization, probes were removed and the cartridges washed extensively with 6 x SSPET using a fluidics station (Affymetrix). Following hybridization, the solutions were removed, the arrays were washed with 6x SSPE-T at 22°C for 7 min, and then washed with 0.5 x SSPE-T at 40°C for 15 minutes. When biotin-labeled RNA was used, the hybridized RNA was stained with a streptavidin-phycoerythrin conjugate (Molecule Probes, Eugene, OR) prior to reading. Hybridized arrays were stained with 2 μg/ml streptavidin-phycoerythrin in 6x SSPE-T at 40°C for 5 minutes. The arrays were read using a scanning confocal microscope made for Affymetrix by Molecular Dynamics (commercially available through Affymetrix, Santa Clara, CA). The scanner uses an argon ion laser as the excitation source, with the emission detected by a photomultiplier tube through either a 530 nm bandpass filter (fluorescein), or a 560 nm longpass filter (phycoerythrin). Nucleic acids of either sense or antisense orientations were used in hybridization

experiments. Arrays with probes for either orientation (reverse complements of each other) are made using the same set of photolithographic masks by reversing the order of the photochemical steps and incorporating the complementary nucleotide.

E. Quantitative Analysis of Hybridization Patterns and Insensitivities

[0244] Following a quantitative scan of an array, or biochip, a grid is aligned to the image using the known dimensions of the array and the corner control regions as markers. The image is reduced to a simple text file containing position and intensity information using software developed at Affymetrix (GENECHIP 3.0 software). This information is merged with another text file that contains information relating physical position on the array to probe sequence and the identity of the RNA and the specific part of the RNA for which the oligonucleotide probe is designed. The quantitative analysis of the hybridization results involves a simple form of pattern recognition based on the assumption that, in the presence of a specific RNA, the PM probes will hybridize more strongly on average than their MM partners. The number of instances in which the PM hybridization signal is larger than the MM signal is computed along with the average of the logarithm of the PM/MM ratios for each probe set. These values are used to make a decision (using a predefined decision matrix) concerning the presence or absence of an RNA. To determine the quantitative RNA abundance, the average of the differences (PM minus MM) for each probe family is calculated. The advantage of the difference method is that signals from random cross-hybridization contribute equally, on average, to

the PM and MM probes, while specific hybridization contributes more to the PM probes. By averaging the pairwise differences, the real signals add constructively while the contributions from cross-hybridization tend to cancel. When assessing the differences between two different RNA samples, the hybridization signals from side-by-side experiments on identically synthesized arrays are compared directly. The magnitude of the changes in the average of the difference (PM-MM) values is interpreted by comparison with the results of spiking experiments as well as the signals observed for the internal standard bacterial and phage RNAs spiked into each sample at a known amount. Data analysis programs developed at Affymetrix, such as the GENECHIP 3.0 software, perform these operation automatically.

[0245] Over 5000 genes were expressed in the central nervous system of the murine samples. Nonetheless, distinct gene expression patterns emerged between peak disease and onset or recovery of the disease. In order to identify the most active genes in EAE development, genes were sought which revealed a pattern of increased or decreased regulation at peak, as compared to onset or recovery. The onset stage served as the baseline comparison, with the assumption that the onset stage was similar in gene expression to undifferentiated naive T cells (i.e., similar to normal control samples samples). The genes demonstrating at least a two-fold increase at peak in EAE tissue as opposed to non-involved tissue are set forth in Table 1, whereas genes having a decrease in expression in EAE tissue as opposed to non-involved tissue are set forth in Table 2. There are several genes which were previously known to be associated with EAE (e.g., which are differently

expressed in MS tissue): TNF-α and IFN-β. The up-regulation of these genes served as validation to the method.

[0246] To generate the data in Tables 1 and 2, genes were clustered hierarchically into groups on the basis of similarity of their expression profiles by the procedure of Eisen et al. ((1998) Proc. Nat'l Acad. Sci. USA 95:25: 14863-8). Genes that were designated absent (A) in all samples in a given experiment were eliminated from the analysis, as were -fold changes over the designated baseline of less than 2. Genes that were present (P) and were of unclear expression (M) were also indicated. Average fold changes between peak/onset and peak/recovery stages were also calculated.

F. Identification of Immune Mediated Genes by Comparison to *In Vitro* Expression Patterns

[0247] In addition, to identify disease-related, antigen-driven immune mediated genes, additional mice were immunized with an EAE-inducing factor and tissue harvested for splenocytes. 3 SJL mice were immunized on Day 0 with 400 μg pertussis toxin and 200 μg injection of myelin oligodendrocyte glycoprotein (MOG) peptide residue 35-55 which, like PLP, induces EAE in mice. On Day 3, an additional dose of 400 μg pertussis toxin was administered. After 10 days, the spleens were harvested, and splenocyte cells were isolated and cultured *in vitro* with 5 μg/ml MOG for 6 and 24 hours, to monitor expression of an immune response at an earlier time (6 hours) and at a later time (24 hours). In addition, a subset of the splenocyte cells were

further treated with 5 μg/ml of anti-IL12p40 (C17.8) antibody, which is believed to produce inhibitory effects on the expression of MS (IL12p40 is an interleukin that induces IFN-γ expression in the Th-1 pathway).

[0248] Isolation of mRNA was performed as described above, with the *in vitro* samples being compared to murine 11,000 genes (Microarray, Affymetrix, Cat Nos. SubA #510243, SubB #510244). As shown in Table 3, MOG restimulation induced differential expression of 199 genes. Of these genes, Tables 4 and 5 indicate 17 genes which further responded to treatment with anti-IL12p40. In addition, quantitative analysis was performed by comparison of the *in vitro* samples to the *in vivo* studies, to identify common genes which were increased or decreased substantially, and/or which were affected by addition by MOG + antiIL12p40. Of the 1265 genes which were found to increase or decrease greater than three fold following MOG restimulation, 184 were found to be in common with genes expressed at peak disease of EAE. Further analysis revealed 6 genes in particular which also responded to treatment with anti-IL12p40 (shown in Table 5): CAPN12, MT1, MYO1F, TLN, UNK_AA117532 and UNK_AA645990.

EXAMPLE 2: IDENTIFICATION AND CHARACTERIZATION OF MARKER cDNA IN HUMAN MODEL OF MULTIPLE SCLEROSIS.

A. Study Design and Subject Entry Criteria

[0249] 60 patients diagnosed with varying stages of multiple sclerosis participated in the study from three sites: University Rochester (Rochester, NY), Lehigh Valley Hospital (Allentown, PA) and Institute of

Neurology, University of London (London, UK). The University of Rochester and Lehigh Valley Hospital provided peripheral blood mononuclear cell (PBMC) samples, while the Institute of Neurology provided brain lesion and non-lesional samples.

[0250] As is characteristic of MS, the patients manifested other neurological disorders and a wide panoply of symptoms, and were at different stages of the disease: relapsing-remitting, 2° progressive, 1° progressive and acute exacerbation. The samples included post-mortem lesional and nonlesional brain samples, as well as other normal controls.

[0251] In general, the donor profile for PBMC samples had the following requirements: no current use of steroids (washout of 30 days), no current use of interferon or copaxone (washout 90 days), no previous history of cyclophosphamide use, no active viral illnesses or infections and no current use of any other study drug.

B. Identification of MS markers from Blood or Brain

[0252] PBMC were isolated from blood samples using Lymphoprep (Nycomed, Oslo, Norway). Lesion and nonlesion brain samples were removed from brain stems and washed To isolate infiltrating immune cells and microglial cells, the samples were run over percoll gradient to remove neurons, endothelial cells and astrocytes.

[0253] As was performed in Experiment 1(C) above, RNA was isolated from the PBMC and brain samples using the

RNeasy mini kit (Quiagen), the RNA probes were hybridized against human chip arrays having approximately 14,000 genes (MicroArray, Affymetrix, cat no. 510448) and quantitative analysis was performed using GeneChip. PBMC samples from MS patients were compared to PBMCs from undiseased pateints, resulting in over 300 differentially regulated genes by more than 2 fold ($p<.01$), as shown in Table 7. Lesions from the MS brain were compared to nonlesion samples, resulting in over 100 genes, as shown in Table 8, that increased or decreased more than 2 fold ($p<.01$). A comparison of the genes expressed in brain and PBMC samples indicated that only 181 genes were shown to be present in both.

C. Comparison to *In Vitro* Expression Patterns

[0254] These genes were compared to genes identified in the murine samples of Example 1(F). Of the 181 genes which were either up- or down-regulated in the human PBMC and brain samples, 6 genes were shown to be in common with the murine genes that were differentially regulated *in vitro*: EEF1D, PIM2, PRDX2, SEC24C, UNK_AJ24 AND XIP (these genes and their accession numbers are listed in Table 9).

D. Taqman Polymerase Chain Reaction

[0255] To ensure that the data obtained from the GeneChip analysis (described above) was reflective of the actual level of gene expression in the cell samples, the expression of 4 selected cytokines, IL-10, IL-8, IL-12p35 and IL-1β, were also measured by polymerase chain reaction, and the results compared to the expression levels of the same cDNAs of the cytokines in non-involved samples and in GeneChip

Table 3

UNK_AA013783	Msa.5720.0	8.5	45	22	0.49	ESTs, Highly similar to INTERFERON-INDUCED 35 KD PROTEIN [Homo sapiens]
DAXX	AF006040	26	68	33	0.49	Fas death domain-associated protein
CASP8	Msa.28381.0	5	17	8	0.47	caspase 8
FAS	FAS	6	17	8	0.47	FAS control sequence (M. musculus) [AFFX]
UNK_ET61797	ET61797	6	13	6	0.46	Mus musculus anti-DNA immunoglobulin heavy chain IgG mRNA, antibody 363s.62, partial cds.
UNK_X52622	Msa.2275.0	4	9	4	0.44	Mouse IN gene for the integrase of an endogenous retrovirus
UNK_ET62928	ET62928	9	16	7	0.44	M.musculus antibody heavy chain variable region (366bp).
HDC	X57437	14.5	55	24	0.44	histidine decarboxylase cluster
ISG15	x56602	83	381	163	0.43	interferon-stimulated protein (15 kDa)
UNK_AA163967	Msa.41726.0	4.5	12	5	0.42	mr24g07.r1 Mus musculus cDNA, 5' end
UNK_AA072573	Msa.28697.0	5.5	46	19	0.41	pre-B-cell colony-enhancing factor
BACTIN5_ST_Mm_A	AFFX-b-ActinMur/M1	7.5	17	7	0.41	Beta-actin 5' control sequence (M. musculus) [AFFX]
TNF	x02611	8	42	17	0.40	tumor necrosis factor
CDKN1A	Msa.1362.0	7.5	23	9	0.39	cyclin-dependent kinase inhibitor 1A (P21)
UBL1A2-PENDING	Msa.20430.0	10.5	26	10	0.38	ubiquitin-like 1 (sentrin) activating enzyme subunit 2
IFIT2	U43085	15.5	94	36	0.38	interferon-induced protein with tetratricopeptide repeats 2
UNK_L32973	L32973	33	130	48	0.37	thymidylate kinase homologue
USP18	aa286391	21.5	106	39	0.37	ubiquitin specific protease 18
BACTIN5_ST_Mm_A	AFFX-b-ActinMur/M1	6.5	14	5	0.36	Beta-actin 5' control sequence (M. musculus) [AFFX]
ACTX	Msa.32904.0	47	174	62	0.36	melanoma X-actin
SCYA4	m35590	4.5	17	6	0.35	small inducible cytokine A4
L-CCR	Msa.22422.0	13	111	37	0.33	lipopolysaccharide inducible C-C chemokine receptor related
RGS1	Msa.37566.0	5	12	4	0.33	regulator of G-protein signaling 1
MAPK13	Msa.6386.0	5	12	4	0.33	mitogen activated protein kinase 13
SCYB10	m33266	14	265	78	0.29	small inducible cytokine B subfamily (Cys-X-Cys), member 10
TNF	TNFa	5	24	7	0.29	TNF control sequence (M. musculus) [AFFX]
IFIT3	I32974	43.5	182	53	0.29	interferon-induced protein with tetratricopeptide repeats 3
MX1	m12279	7	36	10	0.28	myxovirus (influenza virus) resistance 1
ICAM1	x52264	5.5	44	12	0.27	intercellular adhesion molecule
TNF	TNFa	4.5	30	8	0.27	TNF control sequence (M. musculus) [AFFX]
UNK_ET61450	ET61450	4	15	4	0.27	Mus musculus interleukin-2 (IL-2) mRNA, strain-specific RF allele, partial cds.
IFIT1	U43084	27	131	32	0.24	interferon-induced protein with tetratricopeptide repeats 1
UNK_X73040	IL2	4	19	4	0.21	IL2 control sequence (M. musculus) [AFFX]
L-CCR	AF030185	5	24	5	0.21	lipopolysaccharide inducible C-C chemokine receptor related
UNK_ET63426	ET63426	5	29	6	0.21	Mouse strain NOD tumor necrosis factor alpha (TNFA) gene, exon 2 and 3.
IFIT1	aa616578	35.5	179	37	0.21	interferon-induced protein with tetratricopeptide repeats 1
UNK_X73040	IL2	4	23	4	0.17	IL2 control sequence (M. musculus) [AFFX]
MAD	x83106	9	78	9	0.12	Max dimerization protein
MAPKAPK2	Msa.2435.0	8.5	51	5	0.10	MAP kinase-activated protein kinase 2
UNK_D19407	D19407	4	42	4	0.10	MUSGS00778 Mouse 3'-directed cDNA; MUSGS00778; clone mb0657.

Table 4

MOG Induced and Blocked by anti-IL12

Name	Accession No.	Baseline	MOG	WT+MOG+antiIL12p40
D2WSU58E	aa409818	9	7	5
TCFL1	Msa.6010.0	12	17	14
BPGM	X13586	75	62	82
UNK_W10325	w10325	10	7	9
MRJ-PENDING	aa499641	12	12	13
ANXA4	Msa.10010.0	4	6	5
TAF2E	D49439	5	8	4
UNK_AA408555	aa408555	9	6	4
SCD2	w41032	5	6	6
UBCE7	w78245	16	17	13
UNK_AA254104	aa254104	15	4	5
UNK_ET61871	ET61871	7	16	5
UNK_AA472320	aa472320	17	16	15
UNK_AA213318	aa213318	11	11	11
UNK_ET61737	ET61737	20	21	22
UNK_AA198324	AA198324	13	17	13
SDF2	D50646	17	10	11

Table 5

Accession No.	Name	Baseline	MOG	WT+MOG+antiIL12p40	
w64452	CAPN12	5	11	5	calpain 12
V00835	MT1	4	12	4	metallothionein 1
Msa.3234.0	MYO1F	5	16	8	myosin If
x56123-2	TLN	4	16	8	talin
aa645990	UNK_AA645990	4	19	7	ESTs, Highly similar to GLUTAMINYL-TRNA SYNTHETASE [Homo sapiens]
aa117532	UNK_AA117532	7	41	16	ESTs, Moderately similar to HYPOTHETICAL 16.7 KD PROTEIN MRP17-MET14 INTERGENIC REGION [Saccharomyces cerev

Table 6

	A	B	C	D	E
1	Table of Genes Known to be Associated with MS				
2					
3	IL-12p40				
4	IL-10				
5	ICAM-2				
6	GAP-43				
7	a-2 chimmaerin				
8	myelin-transcription factor-1				
9	activin type II receptor				
10	retinoic acid receptor α-1				
11	EF-1				
12	neuron-specific enolase				
13	proton ATPase				
14	Glutaminyl-tRNA synthetase				
15	Na, K-ATPase beta subunit				
16	neuroendocrine-specific protein A				
17	alpha-N-acetylgalactosaminidase				
18	liver type alkaline phosphatase				
19	male-enhanced antigen				
20	mitochondrial enoyl-CoA hydratase				
21	import precursor-ATP synthase α				
22	serine dehydratase				
23	adenine nucleotide translocator-2				
24	ornithine decoboxylase antienzyme				
25	protein kinase C type β-1				
26	cell surface glycoprotein A15				
27	IRF-2				
28	56-K autoantigen annexin XI				
29	high endothelial venule				
30	leukotriene A-4 hydrolase				
31	transcription factor E2A				
32	TNFα receptor-2				
33	duffy chemokine receptor				
34	MS-1				
35	PRG-2				
36	MBP				
37	protein tyrosine phosphatase, receptor type, C				
38	MOG				
39	Interferon β1				
40	ICAM-1				
41	CTLA4				
42	APOE				

Table 7

Gene Name	Accession #	MS Avg	MS SD	Norm Avg	Norm SD	p value	FC MS/Nor	Description
XAP4	U67322	53.20	18.65	3.77	3.03	0.0039	14.11	HBV associated factor
OA48-18	AF089250	52.60	18.97	6.46	4.82	0.0050	8.14	acid-inducible phosphoprotein
KIAA0313	AB002311	13.00	5.34	2.08	1.38	0.0095	6.26	PDZ domain containing guanine nucleotide exchange factor(GEF)1; RA(Ras/Rap1A-associating)-GEF
PPP3CC	S48622	68.60	14.36	12.00	7.72	0.0004	5.72	protein phosphatase 3 (formerly 2B), catalytic subunit, gamma isoform (calcineurin A gamma)
KIAA0742	AB018285	81.80	18.99	15.54	8.07	0.0010	5.26	KIAA0742 protein
UNK_AL022398	AL022398	21.00	5.05	5.85	4.16	0.0009	3.59	Cluster Incl AL022398: Homo sapiens DNA sequence from PAC 434O14 on chromosome 1q32.3-41. Contains the HSD11B1 gene for Hydroxysteroid (11-beta
LSFR2	R56605	19.00	5.52	5.54	5.06	0.0023	3.43	linked to Surfeit genes in Fugu rubripes 2; LSFR2 gene 2
UNK_AI670788	AI670788	19.80	5.72	5.85	4.54	0.0027	3.39	Homo sapiens cDNA FLJ13967 fis, clone Y79AA1001402, weakly similar to Homo sapiens paraneoplastic cancer-testis-brain antigen (MA4) mRNA
TBCC	U61234	16.20	4.87	4.85	3.51	0.0036	3.34	tubulin-specific chaperone c
M6A	AF014837	24.80	9.07	7.46	8.69	0.0078	3.32	putative methyltransferase
BCR	X02596	18.00	4.69	5.85	4.79	0.0015	3.08	breakpoint cluster region
RAB27A	U57094	32.40	4.83	10.54	6.85	0.0000	3.07	RAB27A, member RAS oncogene family
101F6	AF040704	13.40	4.22	4.38	3.12	0.0053	3.06	putative tumor suppressor
PAP	X76770	20.40	6.62	6.77	9.36	0.0057	3.01	poly(A) polymerase
HNRPH3	AF052131	23.80	4.97	7.92	6.37	0.0003	3.00	heterogeneous nuclear ribonucleoprotein H3 (2H9)
DMP1	AI656421	32.60	6.19	10.92	8.99	0.0001	2.98	cyclin D-binding Myb-like protein
LOC56270	AF091083	14.00	2.92	4.69	2.36	0.0006	2.98	hypothetical protein 628
ARCN1	X81198	12.40	2.88	4.23	3.06	0.0008	2.93	archain 1
OGT	U77413	16.40	3.44	5.62	3.28	0.0005	2.92	O-linked N-acetylglucosamine (GlcNAc) transferase (UDP-N-acetylglucosamine:polypeptide-N-acetylglucosaminyl transferase)
GLS	AB020645	74.20	13.83	25.62	16.07	0.0002	2.90	glutaminase
PRKACB	M34181	17.60	4.56	6.08	5.50	0.0015	2.90	protein kinase, cAMP-dependent, catalytic, beta
PTPN22	AF001846	112.60	15.01	39.00	25.96	0.0000	2.89	protein tyrosine phosphatase, non-receptor type 22 (lymphoid)
TRAF5	AB000509	52.00	14.00	18.00	9.84	0.0031	2.89	TNF receptor-associated factor 5
DKFZP564O04	AL050390	65.00	11.25	22.62	12.93	0.0001	2.87	hypothetical protein DKFZp564O043
KIAA0247	D87434	11.00	2.83	3.85	1.99	0.0025	2.86	KIAA0247 gene product
FLT3LG	U03858	36.20	7.76	12.69	12.25	0.0004	2.85	fms-related tyrosine kinase 3 ligand
APBA3	AB021638	48.40	11.41	17.00	9.84	0.0013	2.85	amyloid beta (A4) precursor protein-binding, family A, member 3 (X11-like 2)
KIAA0603	AB011175	16.20	5.17	5.69	3.45	0.0071	2.85	KIAA0603 gene product
KIAA0855	AB020682	18.80	4.21	6.62	4.07	0.0008	2.84	golgin-67
KIAA0540	AB011112	50.00	12.31	17.62	13.91	0.0012	2.84	KIAA0540 protein
HIVEP2	AL023584	12.00	2.45	4.23	2.65	0.0004	2.84	human immunodeficiency virus type I enhancer-binding protein 2
SMC1L1	D80000	27.80	7.40	9.85	8.82	0.0020	2.82	SMC1 (structural maintenance of chromosomes 1, yeast)-like 1
SULT1A2	U34804	11.20	2.28	4.00	2.16	0.0005	2.80	sulfotransferase family, cytosolic, 1A, phenol-preferring, member 2
KIAA0542	AB011114	11.60	3.05	4.15	2.97	0.0022	2.79	KIAA0542 gene product
KIAA0907	AB020714	26.60	7.96	9.54	7.48	0.0045	2.79	KIAA0907 protein
IKBKB	AF031416	21.20	6.72	7.62	5.59	0.0064	2.78	inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase beta
POLR2H	U37689	21.40	6.95	7.69	5.08	0.0077	2.78	polymerase (RNA) II (DNA directed) polypeptide H
MY047	AF063605	28.80	7.19	10.38	6.45	0.0018	2.77	MY047 protein
RNF4	AB000468	61.40	14.33	22.15	12.95	0.0012	2.77	ring finger protein 4
NMT2	AF043325	13.20	4.21	4.77	2.92	0.0074	2.77	N-myristoyltransferase 2
FBXO21	AB020682	10.00	3.08	3.62	1.89	0.0068	2.77	F-box only protein 21
DDX1	X70649	20.40	5.73	7.38	6.02	0.0031	2.76	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 1
HSF1	M64673	15.00	3.08	5.46	3.71	0.0004	2.75	heat shock transcription factor 1
SIP	AF080561	12.20	4.02	4.46	2.40	0.0095	2.73	SYT interacting protein
AQP3	AB001325	13.40	2.41	4.92	3.50	0.0001	2.72	aquaporin 3
MS4A2	M27394	15.00	4.47	5.54	3.99	0.0049	2.71	membrane-spanning 4-domains, subfamily A, member 2
EZH1	AB002386	20.60	4.39	7.62	5.66	0.0005	2.71	enhancer of zeste (Drosophila) homolog 1
RAB2L	AL050259	17.00	1.41	6.31	5.31	0.0000	2.70	RAB2, member RAS oncogene family-like
GRB2	M96995	11.60	1.52	4.31	1.93	0.0000	2.69	growth factor receptor-bound protein 2
PSCD1	M85189	31.20	9.73	11.62	8.33	0.0084	2.69	pleckstrin homology, Sec7 and coiled/coil domains 1(cytohesin 1)
BTN3A3	U90548	18.80	4.82	7.00	6.77	0.0019	2.69	butyrophilin, subfamily 3, member A3
MPV17	X76538	22.00	3.54	8.23	6.78	0.0001	2.67	MpV17 transgene, murine homolog, glomerulosclerosis
TGFBR2	D50683	91.60	26.10	34.31	25.95	0.0038	2.67	transforming growth factor, beta receptor II (70-80kD)
RAF1	X03484	23.20	4.32	8.69	5.79	0.0002	2.67	v-raf-1 murine leukemia viral oncogene homolog 1
KIAA0129	D50919	15.60	4.39	5.85	5.89	0.0035	2.67	KIAA0129 gene product
RBM6	AF069517	24.20	7.29	9.08	6.17	0.0056	2.67	RNA binding motif protein 6
TNRC1	U60760	16.20	3.27	6.08	4.01	0.0004	2.67	trinucleotide repeat containing 1
C3ORF4	AL080097	55.20	8.58	20.77	11.60	0.0000	2.66	chromosome 3 open reading frame 4
ADH5	M81118	11.00	3.24	4.15	2.30	0.0057	2.65	alcohol dehydrogenase 5 (class III), chi polypeptide
UNK_AA44279	AA442799	22.40	7.47	8.46	5.80	0.0095	2.65	Cluster Incl AA442799: zv69b10.r1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone IMAGE:758875 5', mRNA sequence.
NMP200	AJ131186	23.80	4.38	9.00	6.24	0.0002	2.64	nuclear matrix protein NMP200 related to splicing factor PRP19
LANCL1	Y11395	75.40	16.44	28.54	21.94	0.0006	2.64	LanC (bacterial lantibiotic synthetase component C)-like 1
SIPA1	AB005666	39.00	8.03	14.77	8.09	0.0006	2.64	signal-induced proliferation-associated gene 1
UNK_AL02231	AL022312	13.00	2.45	4.92	4.05	0.0002	2.64	Cluster Incl AL022312: Human DNA sequence from clone RP5-1104E15 on chromosome 22q12.3-13.1 Contains the MGAT3 gene for mannosyl (beta-1,4-)-gly
RBM4	U69505	13.20	3.70	5.00	4.47	0.0034	2.64	RNA binding motif protein 4
TUBA3	X01703	13.60	3.13	5.15	3.85	0.0010	2.64	Tubulin, alpha, brain-specific
P130	D21262	16.40	5.03	6.23	4.57	0.0061	2.63	nucleolar phosphoprotein p130

RASGRP2	AI688812	36.00	3.54	13.69	9.99	0.0000	2.63 RAS guanyl releasing protein 2 (calcium and DAG-regulated)
KIAA0746	AB018289	59.20	19.56	22.54	16.44	0.0090	2.63 KIAA0746 protein
HIF1A	U22431	22.00	6.60	8.38	9.33	0.0058	2.62 hypoxia-inducible factor 1, alpha subunit (basic helix-loop-helix transcription factor)
KIAA0906	AB020713	47.80	6.83	18.23	13.98	0.0000	2.62 KIAA0906 protein
RAC1	M29870	80.80	8.70	30.85	18.76	0.0000	2.62 ras-related C3 botulinum toxin substrate 1 (rho family, small GTP binding protein Rac1)
IRF4	U52682	15.20	3.42	5.85	4.54	0.0009	2.60 interferon regulatory factor 4
MTA1	U35113	11.80	3.83	4.54	3.18	0.0087	2.60 metastasis associated 1
PPIE	AF042386	12.20	2.17	4.69	3.20	0.0001	2.60 peptidylprolyl isomerase E (cyclophilin E)
TNRC11	AF071309	22.80	5.40	8.77	7.14	0.0012	2.60 trinucleotide repeat containing 11 (THR-associated protein, 230 kDa subunit)
YME1L1	AJ132637	60.60	15.71	23.54	22.04	0.0024	2.57 YME1 (S.cerevisiae)-like 1
MCM6	D84557	39.80	7.29	15.46	14.32	0.0003	2.57 minichromosome maintenance deficient (mis5, S. pombe) 6
ZNF91	L11672	19.60	6.02	7.62	5.62	0.0065	2.57 zinc finger protein 91 (HPF7, HTF10)
PTD017	AA152202	35.60	10.01	13.85	9.49	0.0041	2.57 PTD017 protein
KIAA0982	AB023199	68.40	21.28	26.62	17.02	0.0075	2.57 KIAA0982 protein
KIAA0909	AB020716	87.80	10.76	34.23	26.47	0.0000	2.56 KIAA0909 protein
SNRPA	M60784	72.20	9.39	28.23	21.55	0.0000	2.56 small nuclear ribonucleoprotein polypeptide A
UNK_AC00438	AC004381	10.80	3.27	4.23	2.98	0.0063	2.55 Cluster Incl AC004381: Homo sapiens Chromosome 16 BAC clone CIT987SK-44M2, complete sequence.
ATP6A1	AA056747	73.20	17.81	28.69	17.72	0.0019	2.55 ATPase, H+ transporting, lysosomal (vacuolar proton pump), alpha polypeptide, 70kD, isoform 1
KIAA0922	AB023139	13.60	3.85	5.38	3.12	0.0050	2.53 KIAA0922 protein
HNRPR	AF000364	40.00	3.81	15.85	11.22	0.0000	2.52 heterogeneous nuclear ribonucleoprotein R
ITM2A	AL021786	12.80	4.15	5.08	3.62	0.0092	2.52 integral membrane protein 2A
TCF3	M31523	79.40	16.92	31.54	24.35	0.0007	2.52 transcription factor 3 (E2A immunoglobulin enhancer binding factors E12/E47)
SOS1	L13858	18.00	3.48	7.15	4.78	0.0003	2.52 son of sevenless (Drosophila) homolog 1
AES	AF072902	41.20	11.28	16.38	12.30	0.0036	2.51 amino-terminal enhancer of split
IGKC	M63438	11.60	3.21	4.62	3.20	0.0040	2.51 immunoglobulin kappa constant
KIAA0099	D43951	23.00	3.61	9.15	7.27	0.0001	2.51 KIAA0099 gene product
HT010	AL049948	15.80	3.27	6.31	4.17	0.0006	2.50 uncharacterized hypothalamus protein HT010
CAST	U31346	10.20	2.39	4.08	2.06	0.0019	2.50 calpastatin
D6S81E	Z37166	10.00	2.55	4.00	3.81	0.0025	2.50 HLA-B associated transcript-1
IL2RG	D11086	10.00	3.08	4.00	2.45	0.0078	2.50 interleukin 2 receptor, gamma (severe combined immunodeficiency)
DKFZP434D15	AA447263	14.80	3.63	5.92	4.41	0.0019	2.50 DKFZP434D156 protein
TYK2	X54637	14.80	2.28	5.92	3.71	0.0000	2.50 tyrosine kinase 2
DUSP11	AF023917	19.40	4.04	7.77	5.48	0.0006	2.50 dual specificity phosphatase 11 (RNA/RNP complex 1-interacting)
ATP2A3	Y15724	32.40	6.47	13.00	7.54	0.0005	2.49 ATPase, Ca++ transporting, ubiquitous
ATR	U49844	13.60	3.13	5.46	4.46	0.0013	2.49 ataxia telangiectasia and Rad3 related
UNK_AF05499	AF054996	26.80	3.19	10.77	6.98	0.0000	2.49 Homo sapiens clone 23763 mRNA sequence
GOSR1	AF047438	13.00	2.92	5.23	3.42	0.0011	2.49 golgi SNAP receptor complex member 1
CPSF1	U37012	50.00	6.44	20.15	13.92	0.0000	2.48 cleavage and polyadenylation specific factor 1, 160kD subunit
DDX19	AJ237946	20.20	4.09	8.15	6.88	0.0006	2.48 DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 19 (Dbp5, yeast, homolog)
WBSCR1	D26068	16.00	4.30	6.46	4.58	0.0035	2.48 Williams-Beuren syndrome chromosome region 1
KIAA0164	D79986	23.80	6.30	9.62	7.82	0.0031	2.48 KIAA0164 gene product
UNK_AI749193	AI749193	27.20	5.63	11.00	9.57	0.0007	2.47 Homo sapiens mRNA; cDNA DKFZp586K2123 (from clone DKFZp586K2123)
SIT	AJ010059	76.60	14.31	31.00	16.03	0.0004	2.47 SHP2 interacting transmembrane adaptor
KIAA0999	AB023216	26.20	8.44	10.62	7.04	0.0097	2.47 KIAA0999 protein
EWSR1	X66899	14.60	4.34	5.92	4.03	0.0063	2.46 Ewing sarcoma breakpoint region 1
SIAT9	AB018356	29.20	5.22	11.85	8.74	0.0002	2.46 sialyltransferase 9 (CMP-NeuAc:lactosylceramide alpha-2,3-sialyltransferase; GM3 synthase)
BRD1	AL080149	14.00	2.45	5.69	2.02	0.0004	2.46 bromodomain-containing 1
JUND	X56681	10.40	1.52	4.23	2.77	0.0000	2.46 jun D proto-oncogene
UNK_AL03128	AL031282	20.80	2.17	8.46	6.60	0.0000	2.46 Cluster Incl AL031282: Human DNA sequence from clone 283E3 on chromosome 1p36.21-36.33. Contains the alternatively spliced gene for Matrix Metalloprot
PURA	M96684	10.20	2.77	4.15	2.34	0.0044	2.46 purine-rich element binding protein A
SRM300	AB002322	17.00	5.24	6.92	4.31	0.0081	2.46 RNA binding protein; AT-rich element binding factor
SMARCA4	U29175	13.40	3.65	5.46	5.55	0.0044	2.45 SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4
DATF1	AB002331	10.00	3.00	4.08	1.98	0.0081	2.45 death associated transcription factor 1
ARHGEF2	U72206	33.20	3.03	13.54	11.01	0.0000	2.45 rho/rac guanine nucleotide exchange factor (GEF) 2
PPP2R5C	Z69030	13.20	3.96	5.38	3.73	0.0068	2.45 protein phosphatase 2, regulatory subunit B (B56), gamma isoform
NCF1	M55067	22.40	3.21	9.15	5.86	0.0000	2.45 neutrophil cytosolic factor 1 (47kD, chronic granulomatous disease, autosomal 1)
UBA52	AC005253	22.00	5.15	9.00	6.35	0.0015	2.44 ubiquitin A-52 residue ribosomal protein fusion product 1
OIP2	AL050353	57.60	13.87	23.62	16.57	0.0018	2.44 Opa-interacting protein 2
PPP2R2A	M64929	24.00	2.92	9.85	6.45	0.0000	2.44 protein phosphatase 2 (formerly 2A), regulatory subunit B (PR 52), alpha isoform
NUMA1	Z11584	11.60	2.17	4.85	3.26	0.0003	2.43 nuclear mitotic apparatus protein 1
SSNA1	Z96932	20.60	3.51	8.46	5.71	0.0001	2.43 Sjogren's syndrome nuclear autoantigen 1
KIAA1096	AL096857	34.80	1.30	14.31	9.89	0.0000	2.43 KIAA1096 protein
EIF5A	U17969	11.40	1.52	4.69	3.28	0.0000	2.43 eukaryotic translation initiation factor 5A
TTC3	D83077	37.00	9.30	15.23	10.88	0.0039	2.43 tetratricopeptide repeat domain 3
STAG2	Z75331	14.20	2.86	5.85	5.13	0.0007	2.43 stromal antigen 2
CC1.3	L10910	19.80	3.70	8.15	7.05	0.0005	2.43 splicing factor (CC1.3)
MYC	V00568	19.80	3.27	8.15	6.50	0.0002	2.43 v-myc avian myelocytomatosis viral oncogene homolog
FLJ20479	AI816034	25.40	8.20	10.46	9.25	0.0099	2.43 likely homolog of yeast Nhp2, component of the H/ACA snoRNP; hypothetical protein FLJ20479
SLC25A1	X96924	24.80	2.17	10.23	5.76	0.0000	2.42 solute carrier family 25 (mitochondrial carrier; citrate transporter), member 1

Table 7

UBE1C	AF046024	11.00	1.58	4.54	2.67	0.0000	2.42 ubiquitin-activating enzyme E1C (homologous to yeast UBA3)
UGP2	U27460	11.00	3.39	4.54	2.30	0.0093	2.42 UDP-glucose pyrophosphorylase 2
HNRPA0	U23803	21.80	6.02	9.00	6.71	0.0043	2.42 heterogeneous nuclear ribonucleoprotein A0
CBF2	M37197	16.00	3.24	6.62	4.27	0.0006	2.42 CCAAT-box-binding transcription factor
RBM5	U23946	58.60	13.52	24.23	16.26	0.0015	2.42 RNA binding motif protein 5
CNIL	AF104398	21.20	4.92	8.77	5.61	0.0018	2.42 cornichon-like
LOC51669	W26659	10.60	1.14	4.38	2.99	0.0000	2.42 HSPC035 protein
KIAA0801	AB018344	63.00	8.60	26.08	16.96	0.0000	2.42 KIAA0801 gene product
SCA2	Y08262	10.40	2.97	4.31	3.01	0.0054	2.41 spinocerebellar ataxia 2 (olivopontocerebellar ataxia 2, autosomal dominant, ataxin 2)
UNK_AL04947	AL049471	10.40	2.70	4.31	1.93	0.0043	2.41 Homo sapiens mRNA; cDNA DKFZp586N012 (from clone DKFZp586N012)
EIF4G1	D12686	20.20	3.90	8.38	6.90	0.0005	2.41 eukaryotic translation initiation factor 4 gamma, 1
POLG	U60325	10.00	2.12	4.15	2.08	0.0011	2.41 polymerase (DNA directed), gamma
GADD45B	AF078077	12.40	2.88	5.15	3.93	0.0016	2.41 growth arrest and DNA-damage-inducible, beta
UNK_X56199	X56199	14.80	2.68	6.15	4.62	0.0003	2.41 Human XIST, coding sequence a mRNA (locus DXS399E)
CSF3R	M59818	12.20	2.49	5.08	3.30	0.0006	2.40 colony stimulating factor 3 receptor (granulocyte)
SLC9A3R1	AF015926	41.00	10.42	17.08	11.42	0.0028	2.40 solute carrier family 9 (sodium/hydrogen exchanger), isoform 3 regulatory factor 1
DKFZP586I222	AL080214	16.40	2.07	6.85	4.41	0.0000	2.40 DKFZP586I2223 protein
ABCC10	AI004207	11.60	2.51	4.85	3.00	0.0010	2.39 ATP-binding cassette, sub-family C (CFTR/MRP), member 10
UNK_AL08021	AL080218	13.80	3.27	5.77	4.53	0.0018	2.39 Homo sapiens mRNA; cDNA DKFZp586N1323 (from clone DKFZp586N1323)
UBE2G2	AF032456	74.60	17.18	31.23	24.26	0.0015	2.39 ubiquitin-conjugating enzyme E2G 2 (homologous to yeast UBC7)
KIAA0546	AB011118	18.00	4.85	7.54	5.16	0.0041	2.39 KIAA0546 protein
PEX10	AA194159	13.40	2.68	5.62	4.31	0.0010	2.39 peroxisome biogenesis factor 10
ADPRT	J03473	17.60	4.83	7.38	6.42	0.0046	2.38 ADP-ribosyltransferase (NAD+; poly (ADP-ribose) polymerase)
LOC51165	Z78324	19.60	3.91	8.23	5.07	0.0006	2.38 SH3 domain-containing protein 6511
HNRPDL	D89678	23.60	3.05	9.92	7.08	0.0000	2.38 heterogeneous nuclear ribonucleoprotein D-like
DKFZP586F13	AL050373	27.60	7.54	11.62	8.10	0.0044	2.38 hypothetical protein
RAD23A	D21235	25.40	2.97	10.69	9.84	0.0002	2.38 RAD23 (S. cerevisiae) homolog A
PPP1CA	S57501	169.20	27.80	71.23	38.15	0.0001	2.38 protein phosphatase 1, catalytic subunit, alpha isoform
RANBP2L1	AF012086	23.00	3.94	9.69	7.59	0.0003	2.37 RAN binding protein 2-like 1
SAFB	U72355	10.40	2.30	4.38	3.73	0.0014	2.37 scaffold attachment factor B
KIAA0082	D43949	67.60	14.88	28.54	18.19	0.0012	2.37 KIAA0082 protein
LOC51580	D14041	18.00	4.80	7.62	5.36	0.0039	2.36 H-2K binding factor-2
PDHB	D90086	21.80	3.70	9.23	8.40	0.0005	2.36 pyruvate dehydrogenase (lipoamide) beta
KIAA0187	D80009	15.80	4.32	6.69	4.52	0.0047	2.36 KIAA0187 gene product
PSMD3	D67025	15.80	1.30	6.69	4.48	0.0000	2.36 proteasome (prosome, macropain) 26S subunit, non-ATPase, 3
DKFZP564M24	AL080119	11.80	3.19	5.00	3.42	0.0044	2.36 DKFZP564M2423 protein
MYCL2	J03069	11.80	2.17	5.00	4.47	0.0006	2.36 v-myc avian myelocytomatosis viral oncogene homolog 2
HERC2	AF041080	31.40	7.70	13.31	9.20	0.0024	2.36 hect domain and RLD 2
UNK_U79273	U79273	13.60	1.14	5.77	4.46	0.0000	2.36 Human clone 23933 mRNA sequence
KIAA1055	AB028978	17.40	3.65	7.38	5.80	0.0009	2.36 KIAA1055 protein
FLJ20259	W27545	15.40	3.78	6.54	4.89	0.0024	2.36 hypothetical protein FLJ20259
KIAA0252	D87440	13.40	3.78	5.69	3.86	0.0056	2.35 KIAA0252 protein
ISG15	AA203213	15.20	4.60	6.46	5.16	0.0080	2.35 interferon-stimulated protein, 15 kDa
DKFZP564A12	W26496	24.60	6.07	10.46	6.01	0.0028	2.35 DKFZP564A122 protein
DKFZP564M24	AL080119	17.00	3.67	7.23	4.59	0.0011	2.35 DKFZP564M2423 protein
PABPC4	U75686	34.00	5.70	14.46	8.72	0.0002	2.35 poly(A)-binding protein, cytoplasmic 4 (inducible form)
SIP2-28	U85611	30.00	2.45	12.77	7.53	0.0000	2.35 calcium and integrin binding protein (DNA-dependent protein kinase interacting protein)
DUT	U31930	13.00	2.74	5.54	4.03	0.0009	2.35 dUTP pyrophosphatase
TIAL1	D64015	22.20	4.92	9.46	5.44	0.0014	2.35 TIA1 cytotoxic granule-associated RNA-binding protein-like 1
DHPS	U26266	20.20	6.02	8.62	4.84	0.0081	2.34 deoxyhypusine synthase
GADD34	U83981	22.00	6.20	9.38	7.35	0.0056	2.34 growth arrest and DNA-damage-inducible 34
KIAA0788	AB018331	11.00	2.35	4.69	3.04	0.0010	2.34 KIAA0788 protein
UNK_AJ22487	AJ224875	11.00	2.92	4.69	3.59	0.0039	2.34 Homo sapiens mRNA for putative glucosyltransferase, partial cds
TMP21	L40397	40.20	5.22	17.15	13.63	0.0001	2.34 transmembrane trafficking protein
MAPK8IP3	AB028989	10.80	2.17	4.62	2.84	0.0006	2.34 mitogen-activated protein kinase 8 interacting protein 3
UNK_AL04995	AL049951	16.20	3.70	6.92	4.27	0.0016	2.34 Homo sapiens mRNA; cDNA DKFZp564O0122 (from clone DKFZp564O0122)
DDX16	AB011149	28.60	8.99	12.23	8.53	0.0099	2.34 DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 16
ERCC5	L20046	10.60	2.07	4.54	2.99	0.0005	2.34 excision repair cross-complementing rodent repair deficiency, complementation group 5 (xeroderma pigmentosum, complementation group G (Cockayne syndr
RALY	L38696	10.60	1.67	4.54	2.76	0.0001	2.34 RNA-binding protein (autoantigenic)
SLC20A1	L20859	13.80	3.49	5.92	4.41	0.0031	2.33 solute carrier family 20 (phosphate transporter), member 1
KIAA1080	AB029003	59.60	3.78	25.62	17.28	0.0000	2.33 KIAA1080 protein; Golgi-associated, gamma-adaptin ear containing, ARF-binding protein 2
NIFU	U47101	13.60	3.65	5.85	4.10	0.0043	2.33 nitrogen fixation cluster-like
CHC1L	AF060219	16.80	3.42	7.23	4.34	0.0008	2.32 chromosome condensation 1-like
FOXO1A	AF032885	26.80	4.92	11.54	8.52	0.0004	2.32 forkhead box O1A (rhabdomyosarcoma)
FSTL1	D89937	53.60	14.84	23.08	21.23	0.0058	2.32 follistatin-like 1
UNK_AL04945	AL049450	13.40	3.85	5.77	3.59	0.0066	2.32 Homo sapiens mRNA; cDNA DKFZp586B1922 (from clone DKFZp586B1922)
ARHGAP1	U02570	16.60	5.13	7.15	5.84	0.0094	2.32 Rho GTPase activating protein 1
IARS	U04953	11.60	2.70	5.00	3.49	0.0019	2.32 isoleucine-tRNA synthetase
PRKCD	D10495	11.60	2.07	5.00	3.58	0.0003	2.32 protein kinase C, delta

Table 7

RNF3	W25793	14.80	3.70	6.38	4.05	0.0030	2.32 ring finger protein 3
TRIM	AJ224878	55.40	15.68	23.92	21.38	0.0068	2.32 T-cell receptor interacting molecule
UNK_N25547	N25547	22.80	5.89	9.85	8.47	0.0039	2.32 Homo sapiens PAK2 mRNA, complete cds
DIPA	U63825	48.80	8.78	21.08	16.76	0.0004	2.32 hepatitis delta antigen-interacting protein A
IGL@	X57809	13.00	2.92	5.62	4.23	0.0015	2.32 immunoglobulin lambda locus
NARS	D84273	122.40	27.81	52.92	47.83	0.0022	2.31 asparaginyl-tRNA synthetase
RBMS1P	D82351	14.40	1.34	6.23	4.38	0.0000	2.31 RNA binding motif, single stranded interacting protein 1, pseudogene
NFATC3	L41067	31.40	6.50	13.62	10.18	0.0009	2.31 nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 3
KNS2	L04733	53.00	11.34	23.00	16.56	0.0011	2.30 kinesin 2 (60-70kD)
RGS2	L13463	23.00	2.45	10.00	6.71	0.0000	2.30 regulator of G-protein signalling 2, 24kD
KIAA0682	AB014582	44.40	10.06	19.31	11.39	0.0017	2.30 KIAA0682 gene product
RING1	Z14000	21.40	3.36	9.31	6.13	0.0001	2.30 ring finger protein 1
PIGH	L19783	12.20	1.30	5.31	3.71	0.0000	2.30 phosphatidylinositol glycan, class H
COPEB	AF001461	15.20	4.15	6.62	5.03	0.0051	2.30 core promoter element binding protein
KIAA0348	AB002344	42.40	9.07	18.46	12.86	0.0011	2.30 KIAA0348 protein
PRKRI	AI095508	16.60	3.85	7.23	4.19	0.0020	2.30 protein-kinase, interferon-inducible double stranded RNA dependent inhibitor
ALOX5AP	AI806222	97.00	23.16	42.31	28.04	0.0023	2.29 arachidonate 5-lipoxygenase-activating protein
FCGR2B	M28696	25.20	6.18	11.00	8.36	0.0028	2.29 Fc fragment of IgG, low affinity IIb, receptor for (CD32)
CD2	M16336	16.20	3.03	7.08	5.82	0.0007	2.29 CD2 antigen (p50), sheep red blood cell receptor
NK4	AA631972	19.00	5.10	8.31	6.32	0.0047	2.29 natural killer cell transcript 4
VAMP2	AL050223	14.60	2.51	6.38	4.35	0.0003	2.29 vesicle-associated membrane protein 2 (synaptobrevin 2)
A2LP	U70671	10.20	2.28	4.46	3.43	0.0017	2.29 ataxin 2 related protein
SMARCA2	D26155	46.40	11.63	20.31	13.36	0.0032	2.28 SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2
Ki-1/57	U77327	21.60	3.21	9.46	5.98	0.0001	2.28 intracellular antigen detected by monoclonal antibody Ki-1; intracellular hyaluronan-binding protein
KIAA0436	AB007898	28.80	6.39	12.54	9.58	0.0017	2.28 putative L-type neutral amino acid transporter
MBD4	AF072250	21.40	5.81	9.38	6.70	0.0051	2.28 methyl-CpG binding domain protein 4
KIAA0737	AB018280	11.40	1.52	5.00	3.85	0.0001	2.28 KIAA0737 gene product
TIF1B	X97548	11.40	2.70	5.00	3.85	0.0024	2.28 KRAB-associated protein 1
CBX1	U35451	21.20	5.26	9.31	6.81	0.0031	2.28 chromobox homolog 1 (Drosophila HP1 beta)
HSPC004	AI827793	46.60	8.76	20.46	14.33	0.0005	2.28 hypothetical protein
HADH2	AF035555	22.40	5.27	9.85	7.87	0.0024	2.28 hydroxyacyl-Coenzyme A dehydrogenase, type II
IL10RA	U00672	12.60	3.05	5.54	4.18	0.0027	2.28 interleukin 10 receptor, alpha
CHD3	U91543	36.20	3.96	15.92	11.35	0.0000	2.27 chromodomain helicase DNA binding protein 3
BUB3	AF047472	26.40	5.13	11.62	8.57	0.0007	2.27 BUB3 (budding uninhibited by benzimidazoles 3, yeast) homolog
NCOR2	U37146	16.80	3.97	7.31	4.53	0.0025	2.27 nuclear receptor co-repressor 2
GILZ	AI635895	13.80	2.05	6.08	4.55	0.0002	2.27 glucocorticoid-induced leucine zipper
POLR2B	X63563	13.80	3.56	6.08	3.59	0.0041	2.27 polymerase (RNA) II (DNA directed) polypeptide B (140kD)
DNM2	L36983	22.00	6.52	9.69	8.22	0.0086	2.27 dynamin 2
P23	L24804	22.00	4.12	9.69	7.85	0.0007	2.27 unactive progesterone receptor, 23 kD
ZYX	X95735	11.00	1.22	4.85	2.97	0.0000	2.27 zyxin
PPP1R2	U68111	30.00	6.89	13.23	9.45	0.0019	2.27 protein phosphatase 1, regulatory (inhibitor) subunit 2
SEC24B	AJ131245	15.00	1.87	6.62	5.20	0.0001	2.27 SEC24 (S. cerevisiae) related gene family, member B
UNK_L29376	L29376	19.00	3.54	8.38	6.56	0.0006	2.27 Homo sapiens (clone 3.8-1) MHC class I mRNA fragment
CEBPB	X52560	17.60	3.21	7.77	5.66	0.0005	2.27 CCAAT/enhancer binding protein (C/EBP), beta
RBL2	X74594	13.40	2.70	5.92	4.11	0.0008	2.26 retinoblastoma-like 2 (p130)
PRPSAP1	D61391	12.00	2.35	5.31	3.73	0.0007	2.26 phosphoribosyl pyrophosphate synthetase-associated protein 1
KIAA0191	D83776	36.00	10.42	15.92	12.19	0.0074	2.26 KIAA0191 protein
NRG1	L12260	21.20	5.07	9.38	5.38	0.0026	2.26 neuregulin 1
POLR2C	J05448	10.60	2.88	4.69	3.12	0.0053	2.26 polymerase (RNA) II (DNA directed) polypeptide C (33kD)
TIEG	S81439	23.80	7.19	10.54	9.02	0.0097	2.26 TGFB inducible early growth response
H-PLK	M55422	13.20	3.49	5.85	5.54	0.0058	2.26 Krueppel-related zinc finger protein
CDC16	U18291	39.40	11.33	17.46	13.25	0.0072	2.26 CDC16 (cell division cycle 16, S. cerevisiae, homolog)
TRA@	M12959	53.60	5.68	23.77	17.62	0.0001	2.26 T cell receptor alpha locus
CEBPD	M83667	10.40	1.52	4.62	3.95	0.0004	2.25 CCAAT/enhancer binding protein (C/EBP), delta
STAT2	U18671	13.00	3.61	5.77	4.46	0.0061	2.25 signal transducer and activator of transcription 2, 113kD
UNK_D28423	D28423	31.20	8.04	13.85	9.61	0.0040	2.25 Cluster Incl D28423: Human mRNA for pre-mRNA splicing factor SRp20, 5'UTR (sequence from the 5'cap to the start codon).
KIAA0594	AB011166	11.60	2.70	5.15	3.26	0.0022	2.25 KIAA0594 protein
UNK_AI659108	AI659108	11.60	2.70	5.15	3.39	0.0022	2.25 Homo sapiens chromosome 19, cosmid R26379
IL17R	U58917	79.80	18.50	35.46	26.69	0.0022	2.25 interleukin 17 receptor
UNK_U53588	U53588	34.60	10.29	15.38	11.05	0.0087	2.25 Cluster Incl U53588: Homo sapiens MHC class 1 region.
NXF1	AJ132712	19.20	3.35	8.54	6.39	0.0004	2.25 nuclear RNA export factor 1
UNK_W73822	W73822	10.20	1.79	4.54	3.31	0.0004	2.25 Cluster Incl W73822: zd52f10.s1 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone IMAGE:344299 3', mRNA sequence.
CLK2	L29218	14.00	2.74	6.23	3.24	0.0007	2.25 CDC-like kinase 2
UNK_AI743654	AI743654	45.80	7.56	20.38	14.13	0.0002	2.25 Homo sapiens cDNA FLJ11585 fis, clone HEMBA1003692
STK17A	AI961743	21.60	5.46	9.62	4.84	0.0041	2.25 serine/threonine kinase 17a (apoptosis-inducing)
KIAA0700	AB014600	26.80	7.44	11.85	7.82	0.0064	2.25 KIAA0700 protein
POU2AF1	Z49194	50.60	6.11	22.54	15.50	0.0000	2.25 POU domain, class 2, associating factor 1
CDC25B	S78187	29.00	3.32	12.92	9.00	0.0000	2.24 cell division cycle 25B
SKIV2L	U09877	29.00	5.83	12.92	8.45	0.0008	2.24 superkiller viralicidic activity 2 (S. cerevisiae homolog)-like

Table 7

KIAA0169	D79991	12.60	2.07	5.62	4.13	0.0003	2.24 KIAA0169 protein
SOD1	X02317	21.40	5.03	9.54	7.10	0.0024	2.24 superoxide dismutase 1, soluble (amyotrophic lateral sclerosis 1 (adult))
STAM	U43899	27.60	5.50	12.31	11.89	0.0021	2.24 signal transducing adaptor molecule (SH3 domain and ITAM motif) 1
HNRPM	L03532	33.80	8.76	15.08	10.24	0.0042	2.24 heterogeneous nuclear ribonucleoprotein M
CDC2L1	M37712	25.00	7.48	11.15	6.35	0.0095	2.24 cell division cycle 2-like 1 (PITSLRE proteins)
GALT	M60091	10.00	2.92	4.46	2.47	0.0085	2.24 galactose-1-phosphate uridylyltransferase
KATNB1	AF052432	21.20	4.27	9.46	6.09	0.0008	2.24 katanin p80 (WD40-containing) subunit B 1
IGHM	X58529	18.20	2.68	7.23	4.85	0.0002	2.24 immunoglobulin heavy constant mu
TRN-SR	AA192359	21.00	5.15	9.38	8.40	0.0031	2.24 transportin-SR
NHP2L1	Z83840	58.00	10.20	25.92	19.73	0.0005	2.24 non-histone chromosome protein 2 (S. cerevisiae)-like 1
FYB	AF001862	11.00	1.87	4.92	2.66	0.0002	2.23 FYN-binding protein (FYB-120/130)
KIAA0174	D79998	11.00	3.32	4.92	3.01	0.0098	2.23 KIAA0174 gene product
UNK_AL03177:	AL031775	11.00	1.87	4.92	2.29	0.0003	2.23 Cluster Incl AL031775: Human DNA sequence from clone 30M3 on chromosome 6p22.1-22.3. Contains three novel genes, one similar to C. elegans Y63D3A.4
ZAP70	L05148	11.00	2.83	4.92	3.48	0.0041	2.23 zeta-chain (TCR) associated protein kinase (70 kD)
CHES1	U58723	13.40	2.51	6.00	4.60	0.0007	2.23 checkpoint suppressor 1
UNK_AA78004	AA780049	13.40	2.51	6.00	3.76	0.0005	2.23 Homo sapiens cDNA: FLJ21439 fis, clone COL04352
KIAA1097	AB029020	52.20	4.76	23.38	13.91	0.0000	2.23 KIAA1097 protein
ATP5A1	D14710	19.40	4.83	8.69	6.09	0.0034	2.23 ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit, isoform 1, cardiac muscle
BAX	U19599	21.80	3.56	9.77	7.13	0.0003	2.23 BCL2-associated X protein
ARRB2	AF106941	10.80	1.48	4.85	2.34	0.0000	2.23 arrestin, beta 2
ATM	U26455	18.00	5.29	8.08	5.85	0.0085	2.23 ataxia telangiectasia mutated (includes complementation groups A, C and D)
CA150	AF017789	14.40	2.70	6.46	4.88	0.0007	2.23 transcription factor CA150
MOF	AI417075	19.20	1.79	8.62	4.17	0.0000	2.23 member of MYST family histone acetyl transferases, homolog of Drosophila MOF
NPIP	AC002045	15.60	4.72	7.00	4.34	0.0100	2.23 nuclear pore complex interacting protein
UNK_AF02681	AF026816	14.40	3.51	6.46	4.70	0.0031	2.23 Homo sapiens putative oncogene protein mRNA, partial cds
MTMR4	AB014547	24.00	3.32	10.77	7.78	0.0001	2.23 myotubularin related protein 4
SF3A3	X81789	13.20	3.11	5.92	4.70	0.0028	2.23 splicing factor 3a, subunit 3, 60kD
PIGC	D85418	43.00	10.68	19.31	10.51	0.0036	2.23 phosphatidylinositol glycan, class C
KIAA0349	AB002347	41.60	5.86	18.69	12.15	0.0001	2.23 KIAA0349 protein
ZNF198	AJ224901	16.60	3.51	7.46	5.17	0.0013	2.22 zinc finger protein 198
TCF8	D15050	15.40	3.44	6.92	4.07	0.0018	2.22 transcription factor 8 (represses interleukin 2 expression)
AP1G2	AI741833	13.00	3.39	5.85	4.24	0.0046	2.22 adaptor-related protein complex 1, gamma 2 subunit
PTPRE	X54134	23.60	6.58	10.62	7.77	0.0066	2.22 protein tyrosine phosphatase, receptor type, E
VHL	L15409	56.60	3.85	25.46	16.16	0.0000	2.22 von Hippel-Lindau syndrome
AP3D1	AF002163	10.60	2.51	4.77	3.72	0.0028	2.22 adaptor-related protein complex 3, delta 1 subunit
HCDI	AI494623	25.80	3.83	11.62	9.20	0.0003	2.22 HCDI protein
MGP	W28510	14.00	2.45	6.31	4.71	0.0005	2.22 matrix Gla protein
IK	AJ005579	60.60	11.15	27.31	21.19	0.0007	2.22 IK cytokine, down-regulator of HLA II
KIAA1028	AB028951	19.80	4.27	8.92	5.68	0.0014	2.22 KIAA1028 protein
LYL1	M22637	32.60	6.11	14.69	7.09	0.0006	2.22 lymphoblastic leukemia derived sequence 1
STK10	AB015718	12.80	2.49	5.77	4.13	0.0008	2.22 serine/threonine kinase 10
DKFZP566H07	AL050060	50.00	10.84	22.54	17.91	0.0018	2.22 DKFZP566H073 protein
KIAA0515	AB011087	18.60	5.41	8.38	5.52	0.0083	2.22 KIAA0515 protein
USP1	AB014458	16.20	4.38	7.31	5.25	0.0056	2.22 ubiquitin specific protease 1
ARIH1	AJ009771	15.00	2.83	6.77	4.78	0.0007	2.22 ariadne (Drosophila) homolog, ubiquitin-conjugating enzyme E2-binding protein, 1
UNK_W26634	W26634	30.00	5.15	13.54	7.21	0.0003	2.22 Cluster Incl W26634: 34b10 Human retina cDNA randomly primed sublibrary Homo sapiens cDNA, mRNA sequence.
ARHGDIA	X69550	48.20	4.09	21.77	17.28	0.0001	2.21 Rho GDP dissociation inhibitor (GDI) alpha
HDAC1	U50079	12.60	3.51	5.69	4.19	0.0067	2.21 histone deacetylase 1
XIP	AF029890	12.60	2.30	5.69	3.86	0.0005	2.21 hepatitis B virus x-interacting protein (9.6kD)
KIAA0365	AB002363	22.80	3.77	10.31	6.51	0.0002	2.21 KIAA0365 gene product
PHF2	AB014562	49.00	13.40	22.15	16.12	0.0061	2.21 PHD finger protein 2
KIAA0544	AB011116	14.80	3.77	6.69	4.44	0.0040	2.21 KIAA0544 protein
ADARB1	U76421	31.80	5.22	14.38	8.53	0.0002	2.21 adenosine deaminase, RNA-specific, B1 (homolog of rat RED1)
MCP	X59405	31.80	6.69	14.38	10.60	0.0014	2.21 membrane cofactor protein (CD46, trophoblast-lymphocyte cross-reactive antigen)
KIAA0117	AA887480	10.20	2.39	4.62	3.12	0.0025	2.21 KIAA0117 protein
SFRS6	AL031681	17.00	1.87	7.69	5.78	0.0001	2.21 splicing factor, arginine/serine-rich 6
KIAA0926	AB023143	26.00	4.64	11.77	8.79	0.0006	2.21 KIAA0926 protein
ITPR3	U01062	15.80	3.03	7.15	4.08	0.0006	2.21 inositol 1,4,5-triphosphate receptor, type 3
EIF5	AL080102	12.40	3.65	5.62	3.20	0.0092	2.21 eukaryotic translation initiation factor 5
NCL	M60858	49.60	6.99	22.46	14.28	0.0001	2.21 nucleolin
RASA1	M23379	14.60	3.21	6.62	4.05	0.0017	2.21 RAS p21 protein activator (GTPase activating protein) 1
PES1	U78310	16.80	3.70	7.62	4.91	0.0017	2.21 pescadillo (zebrafish) homolog 1, containing BRCT domain
RAD21	D38551	24.60	4.39	11.15	6.57	0.0004	2.21 RAD21 (S. pombe) homolog
KIAA1041	AB028964	15.60	2.51	7.08	3.28	0.0002	2.20 KIAA1041 protein
ETR101	M62831	33.40	7.83	15.15	12.03	0.0029	2.20 immediate early protein
E2F4	AA552140	17.80	3.03	8.08	5.44	0.0003	2.20 E2F transcription factor 4, p107/p130-binding
DBP	U48213	22.20	4.55	10.08	11.03	0.0046	2.20 D site of albumin promoter (albumin D-box) binding protein
TEB4	AB011169	14.40	3.58	6.54	3.69	0.0037	2.20 similar to S. cerevisiae SSM4
KIAA1067	AB028990	75.20	12.60	34.15	23.68	0.0003	2.20 KIAA1067 protein

Table 7

SF3B2	U41371	63.00	9.67	28.62	25.14	0.0007	2.20 splicing factor 3b, subunit 2, 145kD
DGSI	L77566	11.00	1.22	5.00	2.16	0.0000	2.20 DiGeorge syndrome critical region gene DGSI
DUSP1	X68277	17.60	1.67	8.00	4.84	0.0000	2.20 dual specificity phosphatase 1
KIAA0911	AB020718	11.00	2.12	5.00	3.79	0.0009	2.20 KIAA0911 protein
RABIF	U74324	27.40	7.96	12.46	9.00	0.0085	2.20 RAB interacting factor
KIAA1111	AB029034	20.80	3.56	9.46	6.79	0.0004	2.20 KIAA1111 protein
MAPRE2	X94232	14.20	3.03	6.46	4.99	0.0017	2.20 microtubule-associated protein, RP/EB family, member 2
GRSF1	U07231	24.00	4.64	10.92	7.94	0.0009	2.20 G-rich RNA sequence binding factor 1
KIAA0752	AB018295	12.00	3.39	5.46	3.99	0.0074	2.20 KIAA0752 protein
PPP2R4	X73478	13.00	2.55	5.92	3.88	0.0008	2.19 protein phosphatase 2A, regulatory subunit B' (PR 53)
CFDP1	AB009265	31.40	5.13	14.31	9.01	0.0002	2.19 craniofacial development protein 1
DED	U51698	31.40	8.14	14.31	9.74	0.0047	2.19 apoptosis antagonizing transcription factor
PSMB7	D38048	23.80	6.50	10.85	6.18	0.0064	2.19 proteasome (prosome, macropain) subunit, beta type, 7
UNK_W32483	W32483	16.20	3.96	7.38	5.64	0.0036	2.19 Human DNA sequence from clone 283E3 on chromosome 1p36.21-36.33. Contains the alternatively spliced gene for Matrix Metalloproteinase in the Female R
CPT1B	Y08683	30.20	7.29	13.77	9.99	0.0032	2.19 carnitine palmitoyltransferase I, muscle
ERPROT213-2	U94836	19.40	4.72	8.85	5.49	0.0032	2.19 protein with polyglutamine repeat; calcium (ca2+) homeostasis endoplasmic reticulum protein
UNK_AL02210	AL022101	19.40	5.50	8.85	6.47	0.0077	2.19 Cluster Incl AL022101: Homo sapiens DNA sequence from PAC 845O24 on chromosome 1p36.1-36.2. Contains a gene for a Heterogenous Nuclear Ribonucle
MADH3	U68019	17.20	3.11	7.85	6.43	0.0009	2.19 MAD (mothers against decapentaplegic, Drosophila) homolog 3
CCT8	D13627	42.80	3.56	19.54	12.86	0.0000	2.19 chaperonin containing TCP1, subunit 8 (theta)
TIM17	X97544	74.80	17.54	34.15	23.54	0.0027	2.19 translocase of inner mitochondrial membrane 17 (yeast) homolog A
DKFZP564O24	AL050015	12.80	3.63	5.85	3.53	0.0078	2.19 DKFZP564O243 protein
GNG10	U31383	12.80	3.03	5.85	4.45	0.0030	2.19 guanine nucleotide binding protein 10
KIAA0073	D38552	22.40	3.13	10.23	7.73	0.0002	2.19 KIAA0073 protein
SFRS10	U68063	12.80	2.39	5.85	3.72	0.0006	2.19 splicing factor, arginine/serine-rich (transformer 2 Drosophila homolog) 10
GATA3	X58072	38.20	6.83	17.46	13.54	0.0007	2.19 GATA-binding protein 3
RRN3	AF001549	28.60	7.02	13.08	9.95	0.0037	2.19 RNA polymerase I transcription factor RRN3
PPIE	AF042385	71.80	12.36	32.85	26.29	0.0007	2.19 peptidylprolyl isomerase E (cyclophilin E)
SATB1	M97287	19.00	4.00	8.69	3.99	0.0016	2.19 special AT-rich sequence binding protein 1 (binds to nuclear matrix/scaffold-associating DNA's)
DKFZP586G01	AL050126	185.60	39.66	84.92	45.01	0.0015	2.19 DKFZP586G011 protein
IMPDH1	J05272	62.20	15.25	28.46	26.51	0.0051	2.19 IMP (inosine monophosphate) dehydrogenase 1
PIG7	AF010312	55.80	13.05	25.54	17.76	0.0027	2.18 LPS-induced TNF-alpha factor
ARHH	Z35227	12.60	2.30	5.77	3.52	0.0005	2.18 ras homolog gene family, member H
DEK	X64229	16.80	3.11	7.69	6.22	0.0010	2.18 DEK oncogene (DNA binding)
HSU53209	U53209	17.80	2.59	8.15	4.69	0.0001	2.18 transformer-2 alpha (htra-2 alpha)
KIAA0554	AB011126	52.20	11.82	23.92	16.21	0.0022	2.18 KIAA0554 protein
BIRC3	U45878	24.00	4.00	11.00	7.87	0.0004	2.18 baculoviral IAP repeat-containing 3
TRB@	X00437	65.60	18.34	30.08	23.73	0.0076	2.18 T cell receptor beta locus
LGALS8	L78132	10.40	2.19	4.77	3.24	0.0014	2.18 lectin, galactoside-binding, soluble, 8 (galectin 8)
SPN	AF049460	26.00	7.52	11.92	7.86	0.0086	2.18 suppressin (nuclear deformed epidermal autoregulatory factor-1 (DEAF-1)-related)
BAF53A	AF041474	43.20	8.61	19.85	12.34	0.0009	2.18 BAF53
NAP1L1	M86667	14.40	2.61	6.62	4.09	0.0005	2.18 nucleosome assembly protein 1-like 1
SSH3BP1	AF006516	44.20	6.76	20.31	16.79	0.0006	2.18 spectrin SH3 domain binding protein 1
FUS1	AF055479	122.20	23.75	56.23	33.99	0.0008	2.17 lung cancer candidate
M11S1	Z48042	11.20	1.64	5.15	2.61	0.0001	2.17 membrane component, chromosome 11, surface marker 1
TRF4	AB005754	11.20	2.28	5.15	3.36	0.0011	2.17 topoisomerase-related function protein 4-1
EVI2B	M60830	24.40	6.23	11.23	7.08	0.0045	2.17 ecotropic viral integration site 2B
PPIE	AF042386	13.20	3.03	6.08	4.54	0.0027	2.17 peptidylprolyl isomerase E (cyclophilin E)
UNK_AI808712	AI808712	14.20	2.28	6.54	4.27	0.0002	2.17 Cluster Incl AI808712: wf57c05.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:2359688 3', mRNA sequence.
AP2B1	M34175	33.40	7.09	15.38	10.67	0.0016	2.17 adaptor-related protein complex 2, beta 1 subunit
PTPRCAP	X97267	33.40	4.16	15.38	8.49	0.0000	2.17 protein tyrosine phosphatase, receptor type, C-associated protein
PSMD9	AI347155	35.40	8.08	16.31	13.73	0.0032	2.17 proteasome (prosome, macropain) 26S subunit, non-ATPase, 9
SFRS2	X75755	20.20	3.96	9.31	5.79	0.0009	2.17 splicing factor, arginine/serine-rich 2
CHKL	AL096780	21.20	5.36	9.77	7.32	0.0045	2.17 choline kinase-like
KIAA0128	AI819942	10.00	1.41	4.62	2.47	0.0001	2.17 KIAA0128 protein; septin 2
SF3A2	L21990	20.00	3.67	9.23	6.73	0.0008	2.17 splicing factor 3a, subunit 2, 66kD
AIM1	U83115	16.00	4.06	7.38	4.65	0.0044	2.17 absent in melanoma 1
CTBP1	U37408	13.00	3.54	6.00	4.93	0.0071	2.17 C-terminal binding protein 1
SYPL	X68194	12.00	1.87	5.54	3.43	0.0002	2.17 synaptophysin-like protein
TMEM4	AB015631	12.00	2.12	5.54	3.71	0.0005	2.17 transmembrane protein 4
UNK_AF07053	AF070535	14.00	2.92	6.46	4.07	0.0013	2.17 Homo sapiens clone 24432 mRNA sequence
KIAA0618	N29685	29.80	4.44	13.77	8.28	0.0001	2.16 KIAA0618 gene product
NRD1	X93209	27.80	3.96	12.85	7.02	0.0001	2.16 nardilysin (N-arginine dibasic convertase)
SNRPF	AI032612	24.80	6.38	11.46	8.53	0.0050	2.16 small nuclear ribonucleoprotein polypeptide F
UNK_AL04112	AL041124	21.80	5.89	10.08	5.79	0.0054	2.16 Cluster Incl AL041124: DKFZp434D0316_s1 434 (synonym: htes3) Homo sapiens cDNA clone DKFZp434D0316 3', mRNA sequence.
NP220	D83032	14.80	2.49	6.85	4.34	0.0003	2.16 nuclear protein
ZNF216	AF062346	14.80	3.27	6.85	4.47	0.0020	2.16 zinc finger protein 216
TNFRSF10B	AF016266	26.60	5.77	12.31	8.68	0.0019	2.16 tumor necrosis factor receptor superfamily, member 10b
ARHGDIB	X69549	35.40	6.27	16.38	11.14	0.0005	2.16 Rho GDP dissociation inhibitor (GDI) beta
BCL6	U00115	11.80	3.03	5.46	3.69	0.0048	2.16 B-cell CLL/lymphoma 6 (zinc finger protein 51)

Table 7

H1FX	D64142	23.60	4.34	10.92	6.55	0.0005	2.16	H1 histone family, member X
SF3B1	AF054284	11.80	2.39	5.46	4.14	0.0014	2.16	splicing factor 3b, subunit 1, 155kD
ETFA	J04058	34.20	9.60	15.85	13.19	0.0085	2.16	electron-transfer-flavoprotein, alpha polypeptide (glutaric aciduria II)
TIC	U63127	24.40	4.72	11.31	7.45	0.0009	2.16	SEC7 homolog
ADAR	X79448	81.00	17.71	37.54	26.80	0.0020	2.16	adenosine deaminase, RNA-specific
TRIP3	L40410	13.60	3.85	6.31	4.91	0.0084	2.16	thyroid hormone receptor interactor 3
LCP2	U20158	11.60	2.19	5.38	3.84	0.0009	2.15	lymphocyte cytosolic protein 2 (SH2 domain-containing leukocyte protein of 76kD)
KIAA0242	D87684	87.80	25.42	40.77	29.33	0.0092	2.15	KIAA0242 protein
HSPA8	Y00371	10.60	2.41	4.92	3.23	0.0024	2.15	heat shock 70kD protein 8
FMNL	AJ008112	25.00	4.30	11.62	9.40	0.0009	2.15	formin-like
LY75	AF011333	86.40	16.68	40.15	27.88	0.0009	2.15	lymphocyte antigen 75
IFITM3	X57352	19.20	5.02	8.92	8.30	0.0075	2.15	interferon induced transmembrane protein 3 (1-8U)
ILF3	U10324	42.20	6.76	19.62	14.49	0.0004	2.15	interleukin enhancer binding factor 3, 90kD
MAP3K4	AF002715	18.20	4.49	8.46	5.81	0.0039	2.15	mitogen-activated protein kinase kinase kinase 4
KIAA0442	AB007902	16.20	2.68	7.54	4.52	0.0003	2.15	KIAA0442 protein
COX11	U79270	14.20	3.63	6.62	4.13	0.0048	2.15	COX11 (yeast) homolog, cytochrome c oxidase assembly protein
RUNX3	Z35278	12.20	2.95	5.69	4.79	0.0045	2.14	runt-related transcription factor 3
UNK_AL05014	AL050148	26.20	5.72	12.23	8.32	0.0020	2.14	Homo sapiens clone 25129 mRNA sequence
UNK_W25986	W25986	16.80	2.59	7.85	5.34	0.0003	2.14	Homo sapiens cDNA FLJ20532 fis, clone KAT10877
SORL1	Y08110	36.40	2.88	17.00	9.33	0.0000	2.14	sortilin-related receptor, L(DLR class) A repeats-containing
KIAA0138	D50928	66.80	19.38	31.23	23.58	0.0098	2.14	KIAA0138 gene product
DIM1	AF023612	10.20	2.39	4.77	2.92	0.0029	2.14	similar to S. pombe dim1+
LOC51317	N98687	10.20	2.28	4.77	3.79	0.0028	2.14	hypothetical protein
TAF2I	X83928	32.40	4.28	15.15	7.70	0.0000	2.14	TATA box binding protein (TBP)-associated factor, RNA polymerase II, I, 28kD
KIAA0766	AB018309	14.80	3.11	6.92	5.56	0.0022	2.14	KIAA0766 gene product
SPTLC1	Y08685	19.40	3.58	9.08	6.44	0.0008	2.14	serine palmitoyltransferase, long chain base subunit 1
VPS26	AF054179	12.00	2.92	5.62	4.35	0.0042	2.14	vacuolar protein sorting 26 (yeast homolog)
LMO4	U24576	33.20	9.28	15.54	10.52	0.0079	2.14	LIM domain only 4
TSC2	X75621	18.40	5.27	8.62	7.29	0.0101	2.14	tuberous sclerosis 2
GPS2	U28963	29.40	7.80	13.77	10.07	0.0062	2.14	G protein pathway suppressor 2
LOC51172	AF052111	40.40	11.44	18.92	11.54	0.0085	2.13	N-acetylglucosamine-1-phosphodiester alpha-N-acetylglucosaminidase
DKFZP566J15	AI857673	26.60	7.50	12.46	10.32	0.0092	2.13	DKFZP566J153 protein
SFRS5	U30826	26.60	5.81	12.46	8.47	0.0020	2.13	splicing factor, arginine/serine-rich 5
FIP2	AF070533	109.60	26.59	51.38	36.86	0.0039	2.13	tumor necrosis factor alpha-inducible cellular protein containing leucine zipper domains; Huntingtin interacting protein L; transcription factor IIIA-interacting prot
EEF1D	Z21507	26.40	4.39	12.38	7.79	0.0003	2.13	eukaryotic translation elongation factor 1 delta (guanine nucleotide exchange protein)
UNK_W27675	W27675	10.00	2.55	4.69	2.46	0.0051	2.13	Cluster Incl W27675: 36b3 Human retina cDNA randomly primed sublibrary Homo sapiens cDNA, mRNA sequence.
LOC55977	AI651368	31.80	5.31	14.92	7.24	0.0003	2.13	hypothetical protein 24636
UNK_AL04930	AL049309	35.40	6.91	16.62	12.09	0.0012	2.13	Homo sapiens mRNA; cDNA DKFZp564B176 (from clone DKFZp564B176)
PPP4C	X70218	37.20	9.63	17.46	12.31	0.0054	2.13	protein phosphatase 4 (formerly X), catalytic subunit
SAS10	AI126004	13.60	2.79	6.38	4.57	0.0016	2.13	disrupter of silencing 10
CD6	X60992	15.40	3.91	7.23	5.64	0.0054	2.13	CD6 antigen
KIAA0100	D43947	55.20	9.88	25.92	20.75	0.0011	2.13	KIAA0100 gene product
RBMS1	X77494	224.20	54.21	105.38	65.37	0.0036	2.13	RNA binding motif, single stranded interacting protein 1
UNK_W72239	W72239	10.80	2.28	5.08	3.28	0.0016	2.13	Homo sapiens mRNA; cDNA DKFZp434M162 (from clone DKFZp434M162)
KIAA1116	AB029039	19.80	5.36	9.31	6.02	0.0068	2.13	KIAA1116 protein
OXA1L	X80695	40.40	8.68	19.00	13.13	0.0019	2.13	oxidase (cytochrome c) assembly 1-like
ASMTL	AA669799	19.60	5.41	9.23	7.07	0.0081	2.12	acetylserotonin O-methyltransferase-like
MDS019	AL078641	22.20	4.60	10.46	7.42	0.0017	2.12	pherbolin-like protein MDS019
CREME9	AF046059	12.40	2.70	5.85	4.74	0.0028	2.12	cytokine receptor-like molecule 9
DAP3	U18321	12.40	3.29	5.85	5.37	0.0085	2.12	death associated protein 3
NME3	U29656	11.40	3.21	5.38	3.43	0.0085	2.12	non-metastatic cells 3, protein expressed in
CKAP1	D49738	32.40	4.04	15.31	10.19	0.0001	2.12	cytoskeleton-associated protein 1
ITGAM	J03925	14.00	2.45	6.62	4.39	0.0006	2.12	integrin, alpha M (complement component receptor 3, alpha; also known as CD11b (p170), macrophage antigen alpha polypeptide)
OAS2	M97434	16.60	3.65	7.85	5.68	0.0024	2.12	2'-5'oligoadenylate synthetase 2
BRF2	U07802	83.00	18.80	39.23	22.22	0.0025	2.12	butyrate response factor 2 (EGF-response factor 2)
USA-CYP	AF016371	117.00	31.44	55.31	37.35	0.0068	2.12	cyclophilin
KIAA0831	AB020638	46.20	8.35	21.85	11.88	0.0005	2.11	KIAA0831 protein
JAK1	M64174	14.80	2.49	7.00	4.12	0.0003	2.11	Janus kinase 1 (a protein tyrosine kinase)
LOC51614	AF091085	14.80	1.92	7.00	5.39	0.0003	2.11	hypothetical 43.2 Kd protein
UNK_AI890903	AI890903	30.40	7.16	14.38	12.80	0.0051	2.11	Cluster Incl AI890903: wm91f10.x1 NCI_CGAP_Ut2 Homo sapiens cDNA clone IMAGE:2443339 3', mRNA sequence.
SMARCE1	AF035262	53.80	7.01	25.46	19.76	0.0004	2.11	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily e, member 1
MLN64	D38255	73.60	19.48	34.85	21.49	0.0062	2.11	steroidogenic acute regulatory protein related
KIAA0579	AB011151	92.60	6.69	43.85	40.19	0.0009	2.11	KIAA0579 protein
UNK_AL04581	AL045811	34.60	9.56	16.38	11.00	0.0079	2.11	Homo sapiens mRNA; cDNA DKFZp434C0814 (from clone DKFZp434C0814)
RNAHP	AF083255	30.20	4.82	14.31	9.61	0.0004	2.11	RNA helicase-related protein
SNRPN	U41303	11.20	3.19	5.31	3.12	0.0093	2.11	small nuclear ribonucleoprotein polypeptide N
KIAA0116	D29958	33.40	5.18	15.85	8.84	0.0002	2.11	KIAA0116 protein
KIAA0663	AB014563	15.40	2.70	7.31	4.92	0.0006	2.11	KIAA0663 gene product
TP53	M22898	15.40	4.04	7.31	5.31	0.0063	2.11	tumor protein p53 (Li-Fraumeni syndrome)

Table 7

BCS1L	AF038195	24.80	3.70	11.77	5.18	0.0001	2.11 BCS1 (yeast homolog)-like
PSMD1	D44466	50.40	11.33	23.92	15.96	0.0026	2.11 proteasome (prosome, macropain) 26S subunit, non-ATPase, 1
OXCT	U62961	16.20	4.15	7.69	5.47	0.0055	2.11 3-oxoacid CoA transferase
DKFZP547E21	AL050261	30.60	6.80	14.54	10.28	0.0026	2.10 DKFZP547E2110 protein
LYPLA1	AF081281	20.40	5.16	9.69	6.07	0.0051	2.10 lysophospholipase I
UNK_L40402	L40402	17.00	3.32	8.08	6.86	0.0022	2.10 Homo sapiens mRNA; cDNA DKFZp434D0935 (from clone DKFZp434D0935)
KIAA1049	AB028972	17.80	2.26	8.46	5.83	0.0002	2.10 KIAA1049 protein
ABLIM	D31883	205.80	34.78	97.85	63.53	0.0005	2.10 actin binding LIM protein 1
HSF2	M65217	11.80	2.59	5.62	3.36	0.0022	2.10 heat shock transcription factor 2
RELA	L19067	11.80	2.39	5.62	3.91	0.0015	2.10 v-rel avian reticuloendotheliosis viral oncogene homolog A (nuclear factor of kappa light polypeptide gene enhancer in B-cells 3 (p65))
NFKB1	M58603	25.20	4.09	12.00	8.02	0.0004	2.10 nuclear factor of kappa light polypeptide gene enhancer in B-cells 1 (p105)
UNK_U50535	U50535	38.60	6.54	18.38	10.87	0.0004	2.10 Human BRCA2 region, mRNA sequence CG006
SULT1A1	X76283	82.20	11.30	39.15	28.38	0.0003	2.10 sulfotransferase family, cytosolic, 1A, phenol-preferring, member 1
DKFZP564O09	W21827	26.00	6.44	12.38	10.17	0.0057	2.10 DKFZP564O092 protein
DKFZP547E10	AL050260	21.80	5.76	10.38	9.40	0.0088	2.10 DKFZP547E1010 protein
UNK_Z82206	Z82206	17.80	1.95	8.38	5.72	0.0001	2.10 Cluster Incl Z82206: Human DNA sequence from clone RP3-370M22 on chromosome 22 Contains part of the gene for ubiquinol-cytochrome c reductase iron-s
CLN3	U32680	26.80	3.63	12.77	8.83	0.0002	2.10 ceroid-lipofuscinosis, neuronal 3, juvenile (Batten, Spielmeyer-Vogt disease)
KIAA0618	AL080109	13.40	2.70	6.38	4.05	0.0013	2.10 KIAA0618 gene product
CD3D	AA919102	64.40	12.64	30.69	20.22	0.0012	2.10 CD3D antigen, delta polypeptide (TIT3 complex)
PCOLN3	U58048	10.00	2.55	4.77	4.40	0.0081	2.10 procollagen (type III) N-endopeptidase
UNK_AJ13311	AJ133115	36.60	9.74	17.46	13.20	0.0072	2.10 Homo sapiens mRNA for TSC-22-like protein
DKFZP727M23	AF055022	11.60	2.61	5.54	4.05	0.0030	2.09 DKFZP727M231 protein
WEE1	W28575	40.60	10.04	19.38	13.52	0.0047	2.09 wee1+ (S. pombe) homolog
FLJ20500	AA522530	17.40	3.91	8.31	5.60	0.0027	2.09 hypothetical protein
CD14	X06882	29.80	6.65	14.23	9.02	0.0025	2.09 CD14 antigen
GOLGA1	U51587	18.20	1.92	8.69	5.48	0.0001	2.09 golgi autoantigen, golgin subfamily a, 1
TIM44	AF026030	30.60	5.94	14.62	9.61	0.0011	2.09 translocase of inner mitochondrial membrane 44 (yeast) homolog
UNK_W26851	W26851	12.40	3.05	5.92	1.89	0.0061	2.09 ESTs, Weakly similar to ALU7_HUMAN ALU SUBFAMILY SQ SEQUENCE CONTAMINATION WARNING ENTRY [H.sapiens]
VPS45B	AJ133421	13.20	3.27	6.31	3.90	0.0046	2.09 vacuolar protein sorting 45B (yeast homolog)
PRKCB1	X07109	40.40	5.55	19.31	12.29	0.0002	2.09 protein kinase C, beta 1
DDX3	U50553	167.20	28.74	79.92	54.70	0.0006	2.09 DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 3
HSA9761	W87858	14.80	3.42	7.08	4.80	0.0032	2.09 putative dimethyladenosine transferase
MATK	S75168	14.80	3.35	7.08	4.91	0.0029	2.09 megakaryocyte-associated tyrosine kinase
KIAA0275	D87465	15.60	2.88	7.46	5.14	0.0009	2.09 KIAA0275 gene product
KIAA0438	AB007898	15.60	3.21	7.46	5.08	0.0017	2.09 KIAA0438 gene product
PRNP	U29185	23.80	2.17	11.38	8.13	0.0000	2.09 prion protein (p27-30) (Creutzfeld-Jakob disease, Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia)
HDAC6	AJ011972	47.40	8.38	22.69	17.73	0.0012	2.09 histone deacetylase 6
GRN	AF055008	21.20	5.89	10.15	7.12	0.0087	2.09 granulin
ITK	L10717	21.20	3.96	10.15	6.58	0.0009	2.09 IL2-inducible T-cell kinase
KIAA0244	D87685	44.80	7.19	21.46	14.91	0.0005	2.09 KIAA0244 protein
UBE2N	D83004	11.40	1.34	5.46	3.89	0.0002	2.09 ubiquitin-conjugating enzyme E2N (homologous to yeast UBC13)
HAX1	U68566	17.00	3.00	8.15	5.49	0.0007	2.08 HS1 binding protein
FCGR2A	M31932	37.20	6.72	17.85	13.27	0.0011	2.08 Fc fragment of IgG, low affinity IIa, receptor for (CD32)
SP2	D28588	18.60	2.07	8.92	6.47	0.0002	2.08 Sp2 transcription factor
CREB1	M27691	103.00	17.76	49.46	29.88	0.0005	2.08 cAMP responsive element binding protein 1
ABR	U01147	30.40	7.89	14.62	10.53	0.0064	2.08 active BCR-related gene
ANPEP	M22324	23.20	4.27	11.15	8.13	0.0011	2.08 alanyl (membrane) aminopeptidase (aminopeptidase N, aminopeptidase M, microsomal aminopeptidase, CD13, p150)
GABBR1	AJ225028	10.40	2.30	5.00	3.08	0.0025	2.08 gamma-aminobutyric acid (GABA) B receptor, 1
KIAA0179	D80001	20.00	4.74	9.62	6.96	0.0041	2.08 KIAA0179 protein
KIAA0630	AB014530	40.80	7.16	19.62	14.74	0.0010	2.08 KIAA0630 protein
SMAP	AW020536	20.80	2.68	10.00	5.42	0.0001	2.08 thyroid hormone receptor coactivating protein
UNK_AL03128	AL031282	54.40	8.53	26.15	17.09	0.0004	2.08 Cluster Incl AL031282: Human DNA sequence from clone 283E3 on chromosome 1p36.21-36.33. Contains the alternatively spliced gene for Matrix Metallo
EED	AF080227	68.20	11.43	32.85	30.82	0.0027	2.08 embryonic ectoderm development
ARHGEF1	Y09160	40.40	3.97	19.46	13.69	0.0001	2.08 Rho guanine nucleotide exchange factor (GEF) 1
PIP5K2B	U85245	18.20	4.76	8.77	6.86	0.0073	2.08 phosphatidylinositol-4-phosphate 5-kinase, type II, beta
UNK_N36842	N36842	18.20	4.66	8.77	6.31	0.0061	2.08 Homo sapiens cDNA FLJ10207 fis, clone HEMBA1005475
SDHA	L21936	17.40	3.21	8.38	5.58	0.0009	2.08 succinate dehydrogenase complex, subunit A, flavoprotein (Fp)
BTG2	U72649	13.40	3.05	6.46	4.16	0.0030	2.07 BTG family, member 2
RPC	Y11651	13.40	3.65	6.46	3.97	0.0079	2.07 RNA 3'-terminal phosphate cyclase
KIAA0135	D50925	64.60	8.65	31.15	22.74	0.0004	2.07 KIAA0135 protein
KIAA0257	D87446	12.60	1.52	6.08	3.30	0.0000	2.07 KIAA0257 protein
UNK_AF00978	AF009787	12.60	2.70	6.08	5.20	0.0038	2.07 Homo sapiens T cell receptor beta chain (TCRB) mRNA, partial cds
HLA-G	M90683	36.20	8.47	17.46	11.84	0.0036	2.07 HLA-G histocompatibility antigen, class I, G
STX7	U77942	11.80	2.49	5.69	3.54	0.0019	2.07 syntaxin 7
MAEA	AF084928	11.00	2.35	5.31	4.35	0.0032	2.07 macrophage erythroblast attacher
SFRS11	M74002	11.00	2.45	5.31	2.36	0.0029	2.07 splicing factor, arginine/serine-rich 11
ARL7	AB016811	10.20	0.84	4.92	3.62	0.0002	2.07 ADP-ribosylation factor-like 7
IGHD	K02882	10.20	2.39	4.92	3.62	0.0040	2.07 immunoglobulin heavy constant delta
TPM4	X05276	10.20	2.39	4.92	3.28	0.0036	2.07 tropomyosin 4

UNK_AI018098	AI018098	256.20	51.21	123.69	70.16	0.0013	2.07 Cluster Incl AI018098: ov65b11.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1642173 3', mRNA sequence.
GG2-1	AF099935	17.20	4.71	8.31	6.21	0.0088	2.07 TNF-induced protein
RAC2	M64595	68.80	11.95	33.23	17.07	0.0005	2.07 ras-related C3 botulinum toxin substrate 2 (rho family, small GTP binding protein Rac2)
SLC2A3	M20681	26.40	4.16	12.77	8.34	0.0004	2.07 solute carrier family 2 (facilitated glucose transporter), member 3
BTN2A1	U90543	19.40	2.07	9.38	6.92	0.0003	2.07 butyrophilin, subfamily 2, member A1
CCT2	AF026166	28.60	5.90	13.85	9.79	0.0020	2.07 chaperonin containing TCP1, subunit 2 (beta)
GLUL	X59834	28.60	7.86	13.85	5.93	0.0094	2.07 glutamate-ammonia ligase (glutamine synthase)
PNN	U77718	17.00	2.92	8.23	5.43	0.0006	2.07 pinin, desmosome associated protein
FASTK	X86779	10.80	1.48	5.23	3.42	0.0002	2.06 Fas-activated serine/threonine kinase
HSGT1	D88208	10.80	2.39	5.23	5.07	0.0068	2.06 suppressor of S. cerevisiae gcr2
LYN	M79321	64.00	14.87	31.00	22.82	0.0040	2.06 v-yes-1 Yamaguchi sarcoma viral related oncogene homolog
ZNF278	AL096890	10.00	1.58	4.85	2.41	0.0002	2.06 zinc finger protein 278
KIAA0370	AB002368	14.60	2.88	7.08	4.42	0.0013	2.06 KIAA0370 protein
NOT56L	Y09022	14.60	2.30	7.08	5.06	0.0006	2.06 Not56 (D. melanogaster)-like protein
UNK_W28732	W28732	115.20	18.19	55.85	26.05	0.0002	2.06 Cluster Incl W28732: 50h7 Human retina cDNA randomly primed sublibrary Homo sapiens cDNA, mRNA sequence.
KIAA1046	AB028959	13.00	2.24	6.31	4.23	0.0007	2.06 KIAA1046 protein
C22ORF4	AL096779	109.80	24.83	53.31	36.41	0.0032	2.06 chromosome 22 open reading frame 4
UNK_AL049415	AL049415	12.20	1.92	5.92	3.75	0.0004	2.06 Homo sapiens mRNA; cDNA DKFZp586N2119 (from clone DKFZp586N2119)
PIG7	AL120815	32.00	5.24	15.54	11.46	0.0008	2.06 LPS-induced TNF-alpha factor
XPOT	AF039022	18.20	2.28	8.85	4.08	0.0000	2.06 exportin, tRNA (nuclear export receptor for tRNAs)
KIAA0978	AB023195	69.60	14.76	33.85	18.58	0.0020	2.06 KIAA0978 protein
UNK_AL022398	AL022398	112.40	25.21	54.69	31.85	0.0028	2.06 Cluster Incl AL022398: Homo sapiens DNA sequence from PAC 434O14 on chromosome 1q32.3.-41. Contains the HSD11B1 gene for Hydroxysteroid (11-beta
37137	Z99716	41.40	8.88	20.15	13.50	0.0006	2.05 septin 3
CD8A	M12824	15.00	3.87	7.31	4.82	0.0049	2.05 CD8 antigen, alpha polypeptide (p32)
KIAA0128	D50918	33.00	5.70	16.08	12.03	0.0011	2.05 KIAA0128 protein; septin 2
LOC64116	AL049963	15.00	4.06	7.31	4.57	0.0081	2.05 up-regulated by BCG-CWS
NFYA	AL031778	54.00	15.22	26.31	18.11	0.0099	2.05 nuclear transcription factor Y, alpha
FLJ20552	AI827877	20.20	5.02	9.85	8.00	0.0087	2.05 hypothetical protein FLJ20552
UNK_AL080111	AL080111	14.20	3.70	6.92	3.20	0.0071	2.05 Homo sapiens mRNA; cDNA DKFZp586G2222 (from clone DKFZp586G2222)
CCM1	U90268	11.20	2.95	5.46	3.76	0.0073	2.05 cerebral cavernous malformations 1
IFNGR2	U05875	43.20	8.98	21.08	14.59	0.0021	2.05 interferon gamma receptor 2 (interferon gamma transducer 1)
SNW1	AF045184	21.60	3.21	10.54	7.14	0.0004	2.05 SKI-INTERACTING PROTEIN
CHRNB1	X14830	10.40	2.61	5.08	3.23	0.0055	2.05 cholinergic receptor, nicotinic, beta polypeptide 1 (muscle)
MAPK9	U09759	10.40	2.79	5.08	3.66	0.0063	2.05 mitogen-activated protein kinase 9
DIA1	M28713	57.60	8.56	28.15	17.47	0.0003	2.05 diaphorase (NADH) (cytochrome b-5 reductase)
ZNFN1A1	U40462	148.40	8.20	72.54	50.50	0.0001	2.05 zinc finger protein, subfamily 1A, 1 (Ikaros)
LMNB2	M94362	50.20	9.63	24.54	17.46	0.0015	2.05 lamin B2
RASGRP1	AF081195	30.20	8.04	14.77	7.79	0.0077	2.04 RAS guanyl releasing protein 1 (calcium and DAG-regulated)
CNP	D13146	18.40	2.97	9.00	4.74	0.0003	2.04 2',3'-cyclic nucleotide 3' phosphodiesterase
KIAA0870	AB020677	82.40	17.07	40.31	29.60	0.0024	2.04 KIAA0870 protein
MAAT1	U19796	31.60	4.16	15.46	10.92	0.0003	2.04 melanoma-associated antigen recognised by cytotoxic T lymphocytes
CACT	Y10319	41.80	7.48	20.46	12.17	0.0007	2.04 carnitine/acylcarnitine translocase
KIAA0253	D87442	13.20	3.56	6.46	4.29	0.0083	2.04 nicastrin
SDHD	AB006202	13.20	3.35	6.46	4.25	0.0060	2.04 succinate dehydrogenase complex, subunit D, integral membrane protein
TASR1	AF047448	11.00	2.45	5.38	3.25	0.0028	2.04 TLS-associated serine-arginine protein 1
UNK_N30151	N30151	11.00	1.41	5.38	3.91	0.0004	2.04 Cluster Incl N30151: yx81f01.s1 Soares melanocyte 2NbHM Homo sapiens cDNA clone IMAGE:268153 3', mRNA sequence.
SFPQ	W27611	38.80	9.20	19.00	12.88	0.0043	2.04 splicing factor proline/glutamine rich (polypyrimidine tract-binding protein-associated)
USP24	AB028980	32.20	7.92	15.77	10.12	0.0051	2.04 ubiquitin specific protease 24
AP3M2	D38293	27.80	3.03	13.62	7.31	0.0000	2.04 adaptor-related protein complex 3, mu 2 subunit
FOS	V01512	27.80	7.73	13.62	9.83	0.0100	2.04 v-fos FBJ murine osteosarcoma viral oncogene homolog
EBI2	L08177	21.20	4.02	10.38	8.45	0.0011	2.04 Epstein-Barr virus induced gene 2 (lymphocyte-specific G protein-coupled receptor)
HSU15552	U15552	14.60	4.04	7.15	4.08	0.0093	2.04 acidic 82 kDa protein mRNA
MYPT1	D87930	14.60	2.70	7.15	5.13	0.0014	2.04 myosin phosphatase, target subunit 1
NCK2	AI275502	14.60	2.07	7.15	4.49	0.0002	2.04 NCK adaptor protein 2
SYNE-1B	AB018339	14.60	3.51	7.15	4.72	0.0048	2.04 synaptic nuclei expressed gene 1b
UNK_AF053356	AF053356	14.60	3.13	7.15	5.89	0.0039	2.04 Cluster Incl AF053356: Homo sapiens chromosome 7q22 sequence, complete sequence.
RMP	AB006572	19.60	1.52	9.62	6.24	0.0001	2.04 RPB5-mediating protein
EIF1A	L18960	15.20	4.15	7.46	5.53	0.0095	2.04 eukaryotic translation initiation factor 1A
ELF4	U32645	18.80	4.97	9.23	6.50	0.0079	2.04 E74-like factor 4 (ets domain transcription factor)
KIAA0397	AB007857	15.80	3.35	7.77	5.56	0.0027	2.03 KIAA0397 gene product
RNASE2	X55988	12.20	3.03	6.00	4.51	0.0063	2.03 ribonuclease, RNase A family, 2 (liver, eosinophil-derived neurotoxin)
NDUFS8	AF038406	41.60	8.44	20.46	13.56	0.0019	2.03 NADH dehydrogenase (ubiquinone) Fe-S protein 8 (23kD) (NADH-coenzyme Q reductase)
DKFZP434M18	AL080141	18.60	4.04	9.15	5.21	0.0025	2.03 DKFZP434M183 protein
76P	AA058852	10.00	2.65	4.92	3.50	0.0081	2.03 gamma tubulin ring complex protein (76p gene)
ARS2	AL096723	20.00	4.74	9.85	4.72	0.0044	2.03 arsenate resistance protein ARS2
C2F	U72514	55.00	11.25	27.08	17.89	0.0020	2.03 putative protein
KIAA0456	AB007925	10.00	2.00	4.92	2.60	0.0015	2.03 KIAA0456 protein
SLU7	AI860856	26.40	8.02	13.00	10.41	0.0049	2.03 step II splicing factor SLU7
HNRPC	M16342	11.40	2.70	5.62	3.15	0.0042	2.03 heterogeneous nuclear ribonucleoprotein C (C1/C2)

Table 7

PRPS1	D00860	11.40	2.88	5.62	3.75	0.0062	2.03 phosphoribosyl pyrophosphate synthetase 1
IL11RA	U32324	24.20	3.83	11.92	6.42	0.0003	2.03 interleukin 11 receptor, alpha
KIAA0276	D87468	12.80	1.46	6.31	3.88	0.0001	2.03 KIAA0276 protein
HNRPD	M94630	49.00	8.37	24.15	14.56	0.0006	2.03 heterogeneous nuclear ribonucleoprotein D (AU-rich element RNA-binding protein 1, 37kD)
RAB1	AL050268	19.80	3.70	9.77	7.49	0.0019	2.03 RAB1, member RAS oncogene family
DJ1042K10.2	AL022238	10.60	2.07	5.23	2.88	0.0013	2.03 hypothetical protein
DDOST	D29643	24.00	6.24	11.85	8.49	0.0076	2.03 dolichyl-diphosphooligosaccharide-protein glycosyltransferase
ERH	D85758	24.00	6.20	11.85	7.34	0.0068	2.03 enhancer of rudimentary (Drosophila) homolog
RPS6KA3	U06316	167.20	31.80	82.54	48.48	0.0011	2.03 ribosomal protein S6 kinase, 90kD, polypeptide 3
DKFZP586O16	AL039458	13.40	3.36	6.62	5.47	0.0079	2.03 DKFZP586O1624 protein
NTE	AJ004832	40.80	4.44	20.15	10.48	0.0000	2.02 neuropathy target esterase
MGEA5	AB014579	26.00	5.10	12.85	7.17	0.0013	2.02 meningioma expressed antigen 5 (hyaluronidase)
UNK_AA17678	AA176780	61.60	15.93	30.46	16.03	0.0069	2.02 Cluster Incl AA176780: zp32a10.s1 Stratagene neuroepithelium (#937231) Homo sapiens cDNA clone IMAGE:611130 3' similar to contains Alu repetitive elem
RERE	AB007927	14.00	3.24	6.92	4.82	0.0042	2.02 arginine-glutamic acid dipeptide (RE) repeats
KIAA1115	AB029038	37.00	8.96	18.31	12.91	0.0016	2.02 KIAA1115 protein
KIAA0154	D63876	163.20	40.83	80.77	56.67	0.0063	2.02 KIAA0154 protein; ADP-ribosylation factor binding protein GGA3
IFI16	M63838	16.00	3.32	7.92	4.68	0.0020	2.02 interferon, gamma-inducible protein 16
RNF15	U90547	16.00	2.45	7.92	5.74	0.0007	2.02 ring finger protein 15
MOF	AL050395	43.80	11.17	21.69	14.66	0.0067	2.02 member of MYST family histone acetyl transferases, homolog of Drosophila MOF
DSIPI	Z50781	55.60	13.59	27.54	20.04	0.0059	2.02 delta sleep inducing peptide, immunoreactor
KIAA0981	AB023198	11.80	2.77	5.85	5.26	0.0078	2.02 KIAA0981 protein
FLJ20297	AF052134	19.40	4.28	9.62	8.11	0.0052	2.02 hypothetical protein FLJ20297
ADAM10	Z48579	15.20	3.27	7.54	5.90	0.0038	2.02 a disintegrin and metalloproteinase domain 10
KIAA0543	AB011115	30.40	7.86	15.08	9.54	0.0071	2.02 KIAA0543 protein
VEGFB	U43368	22.80	2.95	11.31	7.30	0.0002	2.02 vascular endothelial growth factor B
KIAA0747	AB018290	56.60	14.79	28.08	23.28	0.0097	2.02 KIAA0747 protein
UNK_AL04759	AL047596	41.40	7.64	20.54	11.84	0.0009	2.02 Homo sapiens cDNA: FLJ21480 fis, clone COL05034
UNK_AL05037	AL050376	33.80	5.89	16.77	11.40	0.0010	2.02 Homo sapiens mRNA; cDNA DKFZp586J101 (from clone DKFZp586J101)
BIN1	AF001383	29.60	7.09	14.69	9.96	0.0050	2.01 bridging integrator 1
CSPG6	AF020043	23.40	5.59	11.62	8.26	0.0052	2.01 chondroitin sulfate proteoglycan 6 (bamacan)
H326	U06631	17.20	3.70	8.54	5.25	0.0026	2.01 H326
UP	X90858	34.40	7.83	17.08	12.84	0.0044	2.01 uridine phosphorylase
TPST2	AF049891	111.20	13.41	55.23	32.14	0.0001	2.01 tyrosylprotein sulfotransferase 2
CL24751	AF070530	17.80	2.17	8.85	4.62	0.0001	2.01 hypothetical protein, clone 24751
APBB1	L77864	13.00	3.16	6.46	4.41	0.0055	2.01 amyloid beta (A4) precursor protein-binding, family B, member 1 (Fe65)
KHSRP	AA628946	13.00	2.12	6.46	4.86	0.0012	2.01 KH-type splicing regulatory protein (FUSE binding protein 2)
COPS3	AF031647	21.20	3.96	10.54	7.09	0.0014	2.01 COP9 (constitutive photomorphogenic, Arabidopsis, homolog) subunit 3
ROCK1	U43195	24.60	3.85	12.23	7.79	0.0005	2.01 Rho-associated, coiled-coil containing protein kinase 1
BAP1	AF045581	16.40	4.28	8.15	4.43	0.0074	2.01 BRCA1 associated protein-1 (ubiquitin carboxy-terminal hydrolase)
CIRBP	D78134	16.40	4.16	8.15	5.65	0.0069	2.01 cold inducible RNA-binding protein
DKFZP586O02	AL096741	23.20	5.93	11.54	8.54	0.0077	2.01 hypothetical protein
PAPSS1	Y10387	40.20	9.20	20.00	10.69	0.0036	2.01 3-prime-phosphoadenosine 5-prime-phosphosulfate synthase 1
MAF	AF055376	66.60	14.62	33.15	23.10	0.0034	2.01 v-maf musculoaponeurotic fibrosarcoma (avian) oncogene homolog
HLA-DQB1	M16276	48.20	8.53	24.00	20.43	0.0028	2.01 major histocompatibility complex, class II, DQ beta 1
ITPR3	U01062	12.20	1.48	6.08	3.25	0.0001	2.01 inositol 1,4,5-triphosphate receptor, type 3
DYRK1A	D86550	58.20	13.33	29.00	17.54	0.0037	2.01 dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 1A
CD3Z	J04132	31.80	7.82	15.85	10.00	0.0056	2.01 CD3Z antigen, zeta polypeptide (TiT3 complex)
AES	AI951946	23.00	3.46	11.46	4.99	0.0002	2.01 amino-terminal enhancer of split
PTPN1	M31724	28.40	6.35	14.15	9.66	0.0037	2.01 protein tyrosine phosphatase, non-receptor type 1
AMD1	M21154	19.60	4.04	9.77	5.69	0.0020	2.01 S-adenosylmethionine decarboxylase 1
KIAA0810	AB018353	66.20	12.56	33.00	19.00	0.0012	2.01 KIAA0810 protein
KIAA0107	D14663	25.60	6.47	12.77	9.02	0.0070	2.00 KIAA0107 gene product
PRSS15	X76040	155.60	16.32	77.62	46.66	0.0001	2.00 protease, serine, 15
NR3C1	M10901	200.60	28.11	100.08	61.50	0.0003	2.00 nuclear receptor subfamily 3, group C, member 1
RNF6	AJ010346	33.60	5.68	16.77	10.05	0.0006	2.00 ring finger protein (C3H2C3 type) 6
ISG20	U88964	28.20	7.46	14.08	10.44	0.0092	2.00 interferon stimulated gene (20kD)
UNK_U23852	U23852	15.40	2.61	7.69	4.75	0.0007	2.00 Cluster Incl U23852: Human T-lymphocyte specific protein tyrosine kinase p56lck (lck) abberant mRNA, complete cds.
CBFB	L20298	17.40	4.34	8.69	7.30	0.0087	2.00 core-binding factor, beta subunit
ATP5J	AA845575	31.40	7.57	15.69	11.33	0.0058	2.00 ATP synthase, H+ transporting, mitochondrial F0 complex, subunit F6
UNK_AF02326	AF023268	33.40	8.23	16.69	12.11	0.0018	2.00
UNK_U90911	U90911	35.40	9.10	17.69	11.65	0.0073	2.00 Homo sapiens cDNA: FLJ23260 fis, clone COL05804, highly similar to HSU90911 Human clone 23652 mRNA sequence
BM036	AI057607	16.00	4.08	8.00	5.70	0.0074	2.00 uncharacterized bone marrow protein BM036
JMJ	AL021938	16.00	3.67	8.00	5.23	0.0041	2.00 jumonji (mouse) homolog
USP9X	X98296	18.00	2.12	9.00	9.06	0.0045	2.00 ubiquitin specific protease 9, X chromosome (Drosophila fat facets related)

Table 8

Gene Name	Accession #	p value	ave norm/uninv	sd norm/uninv	ave chronic	sd chronic	avg fc chron/uninv	gene description
CCT3	X74801	0.009937	2.75	1.83	35.20	27.31	12.80	CCT3
GRHPR	W28944	0.009949	2.38	1.06	18.40	15.22	7.75	glyoxylate reductase/hydroxypyruvate reductase
ZNF212	U38864	0.008821	3.38	0.92	12.60	6.31	3.73	zinc finger protein 212
RPS3	X55715	0.010192	2.88	2.64	10.60	1.67	3.69	RPS3
SPARC	J03040	0.009001	4.13	4.52	10.00	2.55	2.42	secreted protein, acidic, cysteine-rich (osteonectin)
HLA-DPB1	M83664	0.009253	6.25	3.92	14.20	5.81	2.27	major histocompatibility complex, class II, DP beta 1
RAN	M31469	0.009278	9.50	5.15	21.40	7.89	2.25	RAN, member RAS oncogene family
GSN	X04412	0.00932	6.13	2.59	13.60	5.37	2.22	GSN
KIAA0992	AB023209	0.009415	4.63	1.51	10.20	6.59	2.21	KIAA0992 protein
SEC14L1	D67029	0.009467	12.88	5.72	28.20	10.89	2.19	SEC14 (S. cerevisiae)-like 1
PLTP	L26232	0.009516	6.88	4.36	15.00	5.66	2.18	phospholipid transfer protein
VIM	Z19554	0.009643	19.50	8.50	41.20	9.83	2.11	VIM
RARRES3	AF060228	0.009658	22.75	9.27	47.20	19.46	2.07	retinoic acid receptor responder (tazarotene induced) 3
UNK_L27065	L27065	0.00848	6.00	5.07	12.40	6.58	2.07	Neurofibromatosis 2 Tumor Suppressor
SAH	X80062	0.003295	16.00	8.49	8.00	1.22	0.50	SAH
MADH5	AF010607	0.001577	28.00	15.10	14.00	2.00	0.50	MAD (mothers against decapentaplegic, Drosophila) homolog 5
IGHMBP2	L14754	0.004237	19.63	9.13	9.80	2.17	0.50	immunoglobulin mu binding protein 2
XBP1	Z93930	0.004382	25.25	13.09	12.60	4.04	0.50	X-box binding protein 1
CALCR	L00587	0.004237	51.75	25.04	25.80	8.58	0.50	calcitonin receptor
MDFI	U78313	0.00308	62.63	30.50	31.00	9.92	0.50	MyoD family inhibitor
UNK_AL050030	AL050030	0.003109	13.38	5.18	6.60	2.19	0.49	UNK_AL050030
UNK_W26326	W26326	0.001916	53.88	24.79	26.20	11.03	0.49	Homo sapiens DNA, cosmid clones TN62 and TN82
TUBA3	X01703	0.001219	12.50	5.63	6.00	1.22	0.48	TUBA3
SPUF	AA883101	0.007552	65.25	33.80	30.40	7.96	0.47	secreted protein of unknown function
NFE2L1	L24123	0.006548	28.38	13.29	13.20	2.05	0.47	NFE2L1
COL18A1	AF018081	0.004472	61.38	32.49	28.40	2.51	0.46	collagen, type XVIII, alpha 1
UNK_W27871	W27871	0.004279	42.63	20.56	19.60	2.70	0.46	Cluster Incl W27871: 39a1 Human retina cDNA randomly primed sublibrary Homo sapiens cDNA, mRNA sequence
HSPA2	L26336	0.004581	23.63	10.74	10.80	2.68	0.46	heat shock 70kD protein 2
FRAP1	W29031	0.00772	26.88	13.46	12.00	4.90	0.45	ESTs, Weakly similar to atrophin-1 related protein [H.sapiens]
UNK_AI290480	AI290480	0.007803	26.50	16.35	11.80	3.27	0.45	ESTs, Moderately similar to JC4969 pig-c protein - human [H.sapiens]
KIAA0477	AB007946	0.006066	88.50	47.91	39.20	14.41	0.44	KIAA0477 gene product
MED6	U78082	0.005343	13.13	5.77	5.80	1.79	0.44	RNA polymerase II transcriptional regulation mediator (Med6, S. cerevisiae, homolog of)
SMTN	AI888563	0.005403	24.00	14.08	10.60	3.36	0.44	inositol polyphosphate 5-phosphatase
MLLT4	AB011399	0.004237	150.38	70.56	65.80	15.83	0.44	myeloid/lymphoid or mixed-lineage leukemia (trithorax (Drosophila) homolog); translocated to, 4
LLGL1	AW007029	0.004429	12.38	6.93	5.40	1.14	0.44	lethal giant larvae (Drosophila) homolog 1
UNK_W29012	W29012	0.004233	55.00	24.12	24.00	1.00	0.44	Cluster Incl W29012: 55a6 Human retina cDNA randomly primed sublibrary Homo sapiens cDNA, mRNA sequence
PIGR	X73079	0.003749	87.75	38.30	38.20	8.44	0.44	PIGR
TBL1	AA149637	0.005301	57.25	27.37	24.80	6.22	0.43	transducin (beta)-like 1
KIAA0210	D86965	0.004846	38.88	18.86	16.80	2.77	0.43	KIAA0210 gene product
LTK	D16105	0.004528	19.00	9.53	8.20	2.68	0.43	leukocyte tyrosine kinase
RPS17	M13932	0.004379	14.38	7.39	6.20	1.48	0.43	ribosomal protein S17
GRSF1	U07231	0.004362	28.88	13.57	12.40	2.07	0.43	G-rich RNA sequence binding factor 1
UNK_AL022398	AL022398	0.004054	20.50	8.70	8.80	3.56	0.43	Cluster Incl AL022398: Homo sapiens DNA sequence from PAC 434O14 on chromosome 1q32.3.-41.
GATA2	M77810	0.00394	17.75	10.46	7.60	2.51	0.43	GATA-binding protein 2
PABPC1	Y00345	0.003317	27.63	11.90	11.80	2.77	0.43	PABPC1
MPHOSPH9	N23137	0.002013	10.38	4.93	4.40	1.14	0.42	M-phase phosphoprotein 9
NFKB1	M58603	0.002132	10.88	5.28	4.60	1.14	0.42	nuclear factor of kappa light polypeptide gene enhancer in B-cells 1 (p105)
SNRPG	AI803447	0.00347	59.00	31.72	24.80	7.26	0.42	small nuclear ribonucleoprotein polypeptide G
KRTHA3B	X82634	0.001769	11.50	6.07	4.80	1.48	0.42	KRTHA3B
CLUL1	D63813	0.00245	21.63	10.97	9.00	2.45	0.42	clusterin-like 1 (retinal)
PMS2	U13696	0.001807	46.63	20.07	19.40	3.91	0.42	postmeiotic segregation increased (S. cerevisiae) 2
MGAT2	U15128	0.002266	12.13	6.42	5.00	1.22	0.41	mannosyl (alpha-1,6-)-glycoprotein beta-1,2-N-acetylglucosaminyltransferase
UBE2L6	AA883502	0.003066	15.63	7.82	6.40	1.14	0.41	ubiquitin-conjugating enzyme E2L 6
GYPE	X53004	0.002639	16.13	9.67	6.60	2.07	0.41	GYPE
IRS1	S62539	0.001856	89.38	43.55	36.40	12.56	0.41	insulin receptor substrate 1
ITGB3	M25108	0.001833	10.50	6.78	4.20	0.84	0.40	integrin, beta 3 (platelet glycoprotein IIIa, antigen CD61)
CDK5R1	X80343	0.004472	16.63	7.93	6.20	3.19	0.37	CDK5R1
UNK_AI557322	AI557322	0.004237	17.25	8.58	6.40	2.07	0.37	Cluster Incl AI557322: PT2.1_16_F11.r tumor2 Homo sapiens cDNA 3', mRNA sequence.
LOC55977	AF070843	0.004237	32.38	17.14	12.00	1.73	0.37	LOC55977
HIPK3	AF004849	0.002643	11.38	6.48	4.20	1.30	0.37	homeodomain-interacting protein kinase 3
PTPN21	X79510	0.002454	17.88	10.92	6.60	1.14	0.37	PTPN21

Table 8

UNK_X89059	X89059	5.03E-05	21.13	10.70	7.40	1.14	0.35	STK9
KIAA0744	AB018287	0.00017	18.88	11.47	6.60	2.51	0.35	KIAA0744 gene product; histone deacetylase 7
SLC18A2	L14269	0.001401	23.88	18.39	8.20	1.92	0.34	solute carrier family 18 (vesicular monoamine), member 2
CELSR1	AI275081	0.000276	25.25	13.22	8.40	2.19	0.33	cadherin EGF LAG seven-pass G-type receptor
THPO	L33410	0.001154	46.25	23.27	15.00	3.00	0.32	thrombopoietin (myeloproliferative leukemia virus oncogene ligand, megakaryocyte growth and development factor
KIAA0368	AB002366	0.000906	10.50	6.14	3.40	0.55	0.32	KIAA0368 protein
UNK_W25917	W25917	0.001177	16.75	10.05	5.40	1.67	0.32	Cluster Incl W25917: 14h6 Human retina cDNA randomly primed subllbrary Homo sapiens cDNA, mRNA sequence
CXADR	Y07593	0.000704	10.63	8.86	3.40	1.14	0.32	CXADR
NME2	X58965	0.000614	54.00	28.17	16.40	2.88	0.30	NME2
ELK1	AB016194	0.000592	18.50	9.18	5.60	1.67	0.30	ELK1, member of ETS oncogene family
NACA	AF054187	0.000396	11.25	7.34	3.40	1.14	0.30	nascent-polypeptide-associated complex alpha polypeptide
ATP5G3	U09813	0.000457	21.25	11.36	6.40	1.52	0.30	ATP synthase, H+ transporting, mitochondrial F0 complex, subunit c (subunit 9) isoform 3
ATP7B	U11700	0.000559	14.75	8.38	4.40	0.89	0.30	ATPase, Cu++ transporting, beta polypeptide (Wilson disease)
HIPK3	AJ523538	0.00053	23.50	12.71	7.00	2.45	0.30	homeodomain-interacting protein kinase 3
FLJ10871	U96629	0.000317	12.13	6.56	3.60	0.89	0.30	UNK_U96629
TRHR	D85376	6.29E-05	12.88	6.94	3.80	1.30	0.30	thyrotropin-releasing hormone receptor
DRD3	U25441	9.9E-06	30.50	15.69	9.00	3.67	0.30	dopamine receptor D3
MYO9B	U42391	1.44E-05	10.88	7.28	3.20	1.64	0.29	myosin IXB
HYPB	AL049470	2.26E-05	12.25	6.69	3.60	1.67	0.29	HYPB
CASP8	X98176	0.000349	11.63	6.25	3.40	1.14	0.29	CASP8
RFPL3	AJ010232	6.98E-05	43.13	23.24	12.60	5.13	0.29	RFPL3
SCN1B	L10338	2.89E-05	16.63	9.53	4.80	2.39	0.29	sodium channel, voltage-gated, type I, beta polypeptide
ARC	D87468	3.48E-05	24.25	17.43	7.00	3.32	0.29	activity-regulated cytoskeleton-associated protein
EGR1	X52541	4.09E-05	10.50	6.52	3.00	0.00	0.29	early growth response 1
UNK_M21259	M21259	0.000202	16.25	8.83	4.60	1.52	0.28	small nuclear ribonucleoprotein polypeptide E
DSPG3	U59111	9.61E-05	17.00	9.50	4.80	0.84	0.28	dermatan sulphate proteoglycan 3
ITK	L10717	0.000154	10.13	7.95	2.80	1.30	0.28	IL2-inducible T-cell kinase
NHLH2	M98740	4.18E-05	10.13	5.69	2.80	1.30	0.28	NHLH2
SREC	D86864	9.83E-05	17.38	9.26	4.80	1.48	0.28	acetyl LDL receptor; SREC
ZNF148	AJ236885	0.00039	10.88	9.72	3.00	0.71	0.28	ZNF148
SMARCD1	U66617	0.000118	13.88	7.77	3.80	1.30	0.27	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily d, member 1
NN8-4AG	U50383	0.000194	10.63	6.44	2.60	0.55	0.24	retinoic acid responsive
UNK_M20642	M20642	0.000122	14.00	6.41	3.40	1.52	0.24	Human alkali myosin light chain 3 mRNA, complete cds
GFAP	S40719	0.000161	10.75	6.94	2.60	0.89	0.24	glial fibrillary acidic protein
KIAA0847	AB020654	0.000141	14.38	8.02	3.40	0.89	0.24	KIAA0847 protein

Table 9

Genes Differentially regulated in Murine EAE and Human MS

Gene Name	Murine Accession	Description
EEF1D	aa253918	Cluster Incl Z21507: H.sapiens EF-1delta gene encoding human elongation factor-1-delta.
PIM2	Msa.2067.0	Human pim-2 protooncogene homolog pim-2h mRNA, complete cds.
PRDX2	Z21848	Cluster Incl L19185: Human natural killer cell enhancing factor (NKEFB) mRNA, complete cds.
SEC24C	aa175517	Cluster Incl D38555: Human mRNA for KIAA0079 gene, complete cds.
UNK_AJ2454	u85207	Cluster Incl AJ245416: Homo sapiens mRNA for G7b protein (G7b gene, located in the class III region of the major histocompatibility complex.
XIP	Msa.3679.0	Cluster Incl AF029890: Homo sapiens hepatitis B virus X interacting protein (XIP) mRNA, complete cds.

Table 10

	A	B	C	D	E
1	Peripheral Blood Genes Predicitve of Multiple Sclerosis				
2					
3	Gene	Observed	Random Perm 1%	Random Perm 5%	Randon Perm (Median)
4	FOS	2.562527	1.4817816	1.1614933	0.7007793
5	U19261	2.114882	1.3043897	0.87986296	0.53156734
6	DUSP1	1.939225	1.2410891	0.8299547	0.5107011
7	KIAA0906	1.87203	1.2133218	0.80576986	0.48326102
8	MAP4K1	1.860437	1.2085855	0.79408497	0.47091565
9	DGKZ	1.837408	1.2069645	0.77625513	0.4690797
10	SNRP70	1.81934	1.1756094	0.7738778	0.45766842
11	ETR101	1.801398	1.1731951	0.76643366	0.440602
12	KIAA0864	1.775063	1.1709846	0.7597211	0.43995833
13	SFRS8	1.77253	1.1582581	0.7580478	0.43687758
14	AI890903	1.764722	1.1560028	0.7550534	0.4319513
15	JUNB	1.752881	1.1488212	0.74930257	0.4269654
16	STAT4	1.719926	1.1444192	0.7461597	0.41349095
17	AQP3	1.705897	1.130565	0.7446831	0.4113997
18	HSU79253	1.664187	1.1186187	0.7290762	0.4062087
19	BIN1	1.649725	1.1111517	0.7287055	0.402228
20	ATM	1.641806	1.1030456	0.7274922	0.39897352
21	EDG4	1.634613	1.0987875	0.7267664	0.39316127
22					
23	Prediction based on correlation metrics proposed by Gollub et. al. (1999)				

WHAT IS CLAIMED:

1. A method of diagnosing a subject with multiple sclerosis, the method comprising the step of comparing:

a) a level of expression of a marker in a sample from the subject, wherein the marker is selected from the group consisting of markers listed in Tables 7-10, and

b) a normal level of expression of the marker in a control sample,

wherein a substantial difference between the level of expression of the marker in the sample from the subject and the normal level is an indication that the subject is afflicted with multiple sclerosis.

2. The method of claim 1, wherein the marker corresponds to a transcribed polynucleotide or a portion thereof.

3. The method of claim 2, wherein the sample is collected from brain tissue.

4. The method of claim 2, wherein the sample is peripheral blood mononuclear cells.

5. The method of claim 1, wherein the control sample is from a non-diseased subject and the substantial difference is a factor of at least about 2.

6. The method of claim 1, wherein the control sample is from non-involved tissue of the subject and the substantial difference is a factor of at least about 2.

7. The method of claim 1, wherein the control sample is from non-involved tissue of the subject and the substantial difference is a factor of at least about 5.

8. The method of claim 1, wherein the level of expression of the marker in the sample is assessed by detecting the presence in the sample of a protein corresponding to the marker.

9. The method of claim 8, wherein the presence of the protein is detected using a reagent which specifically binds with the protein.

10. The method of claim 9, wherein the reagent comprises an antibody or fragments thereof.

11. The method of claim 1, wherein the marker is selected from the markers listed in Table 9.

12. The method of claim 1, wherein the marker is selected from the markers listed in Table 10.

13. The method of claim 1, wherein the level of expression of the marker in the sample is assessed by detecting the presence in the sample of a transcribed polynucleotide or portion thereof, wherein the transcribed polynucleotide comprises the marker.

14. The method of claim 13, wherein the transcribed polynucleotide is a mRNA.

15. The method of claim 13, wherein the transcribed polynucleotide is a cDNA.

16. The method of claim 1, wherein the level of expression of the marker in the sample is assessed by detecting the presence in the sample of a transcribed polynucleotide or a portion thereof which hybridizes with a labeled probe under stringent conditions, wherein the transcribed polynucleotide comprises the marker.

17. A method of diagnosing a subject with multiple sclerosis, the method comprising the step of comparing:

a) a level of expression in a sample of the subject of each of a panel of markers independently selected from the markers listed in Tables 7-10, and

b) a normal level of expression of the panel of markers obtained from a control sample,

wherein the level of expression of the panel of markers is substantially different, relative to the corresponding normal level of expression of the panel of markers, indicates that the subject is afflicted with multiple sclerosis.

18. The method of claim 17, wherein the panel of markers comprises at least 5 markers.

19. The method of claim 17, wherein the control sample is from a non-diseased subject.

20. The method of claim 17, wherein the control sample is from non-involved tissue of the subject.

21. The method of claim 17, wherein the panel of markers comprises markers listed in Table 10.

22. A method for monitoring the progression of multiple sclerosis in a subject, the method comprising the steps of:

a) detecting in a subject sample at a first point in time, a level of expression of at least one marker, wherein the at least one marker is selected from the group consisting of markers listed in Tables 7-10;

b) repeating step a) at a subsequent point in time with the at least one marker;

c) detecting a substantial difference between the levels of expression detected in steps a) and b);

wherein the substantial difference between the levels of expression indicates that the subject has progressed to a different stage of multiple sclerosis.

23. The method of claim 22, wherein at least 5 markers are selected from the group of markers listed in Tables 7-10.

24. The method of claim 22, wherein the at least one marker corresponds to a transcribed polynucleotide or portion thereof.

25. The method of claim 24, wherein the samples are collected from brain tissue.

26. The method of claim 24, wherein the samples are peripheral blood mononuclear cells.

27. A method of assessing the efficacy of a test compound for inhibiting multiple sclerosis in a subject, the method comprising the step of comparing:

a) expression of a marker in a first sample obtained from the subject and exposed to the test

compound, wherein the marker is selected from the group consisting of markers listed in Tables 7-10, and

b) expression of the same marker in a second sample obtained from the subject, wherein the second sample is not exposed to the test compound,

wherein a substantially different level of expression of the marker in the first sample, relative to the second sample, is an indication that the test compound is efficacious for inhibiting multiple sclerosis in the subject.

28. The method of claim 27, wherein the first and second samples are portions of a single sample obtained from the subject.

29. The method of claim 28, wherein the level of expression in the first sample approximates the level of expression in a control sample.

30. A method of assessing the efficacy of a therapy for inhibiting multiple sclerosis in a subject, the method comprising the steps of comparing:

a) expression of a marker in a first sample obtained from the subject prior to providing at least a portion of the therapy to the subject, wherein the marker is selected from the group consisting of markers listed in Tables 7-10, and

b) expression of the marker in a second sample following provision of the portion of the therapy,

wherein a substantially different level of expression of the marker in the second sample, relative to the first sample, is an indication that the test compound is efficacious for inhibiting multiple sclerosis in the subject.

31. The method of claim 30, wherein a substantially similar level of expression in the second sample, relative to the control sample, is an additional indication that the test compound is efficacious for inhibiting multiple sclerosis.

32. The method of claim 31, further comparing expression of the marker in a control sample, wherein expression of the marker in the second sample is similar to expression of the marker in the control sample.

33. A method of screening test compounds for inhibitors of multiple sclerosis, the method comprising the steps of:

a) obtaining a sample comprising cells from the subject;

b) separately maintaining aliquots of the sample in the presence of a plurality of test compounds;

c) comparing the expression levels of a marker in each of the aliquots, wherein the marker is selected from the group consisting of markers listed in Tables 7-10; and

d) selecting one of the test compounds which induces a substantially different level of expression of the marker in the aliquot containing that test compound, relative to other test compounds.

34. The method of claim 33, wherein the substantially different level of expression is a substantially lower level of expression.

35. The method of claim 34, wherein the substantially different level of expression is a substantially enhanced level of expression.

36. A kit for diagnosing a subject with multiple sclerosis, the kit comprising reagents for assessing expression of a marker selected from the group consisting of markers listed in Tables 7-10.

37. A kit for diagnosing multiple sclerosis in a subject, the kit comprising a nucleic acid probe wherein the probe specifically binds with a transcribed polynucleotide corresponding to a marker selected from the group consisting of markers listed in Tables 7-10.

38. A kit for assessing the suitability of each of a plurality of compounds for inhibiting progression of multiple sclerosis in a subject, the kit comprising:

a) the plurality of compounds; and

b) a reagent for assessing expression of a marker selected from the group consisting of markers listed in Tables 7-10.

39. A kit for diagnosing multiple sclerosis in a subject, the kit comprising an antibody, wherein the antibody specifically binds with a protein corresponding to a marker selected from the group consisting of markers listed in Tables 7-10.

40. A method of modulating a level of expression of a marker selected from the markers listed in Tables 7-10, the method comprising providing to diseased cells of a subject an antisense oligonucleotide complementary to a polynucleotide corresponding to the marker.

41. A method of modulating a level of expression of a marker selected from the markers listed in Tables 7-10, the method comprising providing to diseased cells of a subject a protein.

42. The method of claim 41, wherein the protein is provided to the cells by providing a vector comprising a polynucleotide encoding the protein to the cells.

43. A method of modulating a level of expression of a marker selected from the markers listed in Tables 7-10, the method comprising providing to diseased cells of a subject an antibody.

44. The method according to claim 43, wherein the method further comprises a therapeutic moiety conjugated to the antibody.

45. A method of localizing a therapeutic moiety to diseased tissue comprising exposing the tissue to an antibody which is specific to a protein encoded from a marker listed in Tables 7-10.

46. A method of screening for a test compound capable of modulating the activity of a protein encoded from a marker listed in Tables 7-10, the method comprising combining the protein and the test compound, and determining the effect of the test compound on the therapeutic efficacy of the protein.

47. A method of screening for a bioactive agent capable of interfering with the binding of a protein encoded from a marker listed in Tables 7-10 and an antibody which binds to the protein, the method comprising:

a) combining the protein, a bioactive agent and an antibody which binds to the protein; and

b) determining the binding of the protein or fragment thereof and the antibody.

48. An antibody which specifically binds to a protein encoded from a marker listed in Tables 7-10.

49. The antibody of claim 48, wherein the antibody is a monoclonal antibody.

50. The antibody of claim 49, wherein the antibody is a humanized antibody.

51. A peptide encoded from markers listed in Tables 7-10.

52. A composition comprising the peptide of claim 51.

53. A composition capable of modulating an immune response in a subject, the composition comprising a protein encoded from a marker listed in Tables 7-10, and a pharmaceutically acceptable carrier.

54. A biochip comprising a panel of markers selected from the group of markers listed in Tables 7-10.

55. The biochip of claim 54, wherein the markers are selected for subjects suspected of having secondary progressive multiple sclerosis.

56. The biochip of claim 54, wherein the markers are selected for subjects suspected of having primary progressive multiple sclerosis.

57. The biochip of claim 54, wherein the markers are selected for subjects suspected of having relaxing-remitting multiple sclerosis.

58. The biochip of claim 54, wherein the markers are selected for subjects suspected of having multiple sclerosis, wherein the subjects are from a high-risk geographic region.

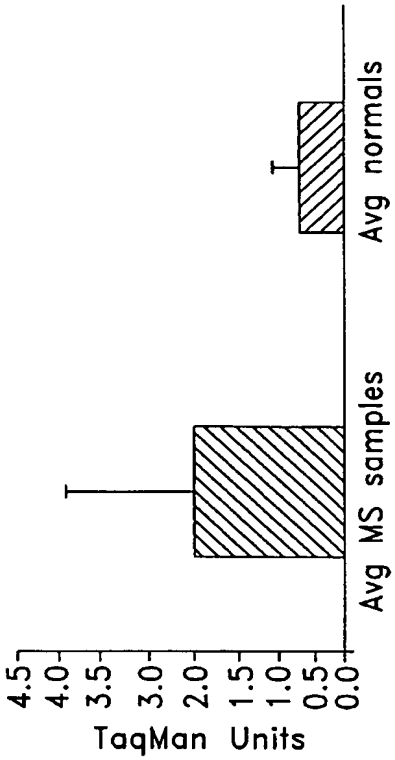

FIG. 1B

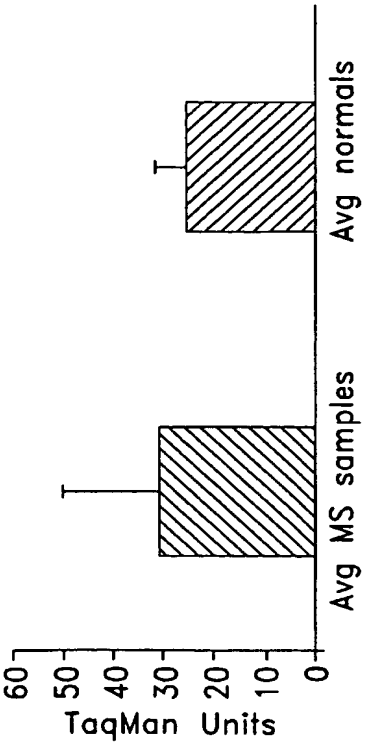

FIG. 1A

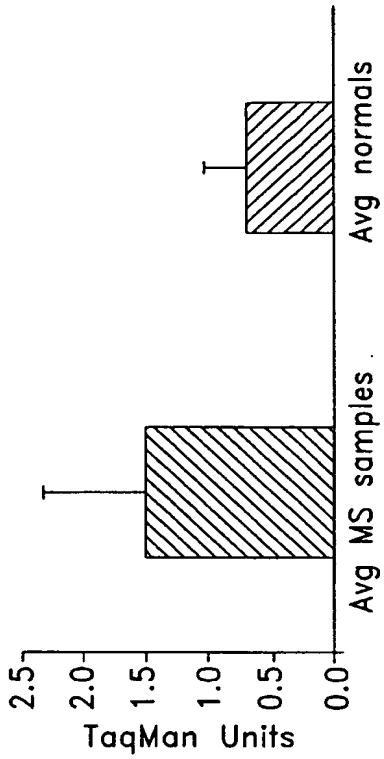

FIG. 1D

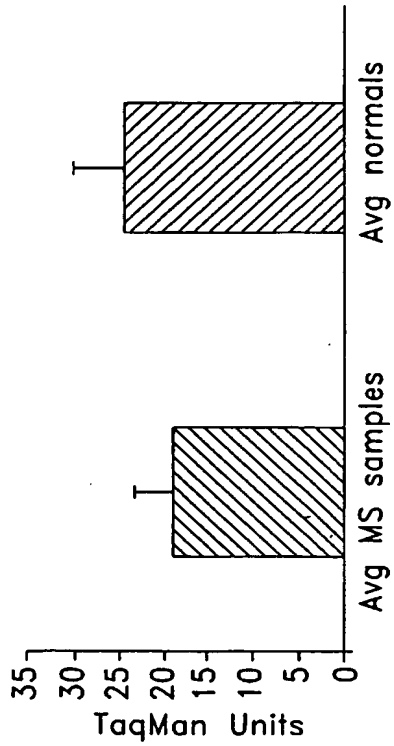

FIG. 1C

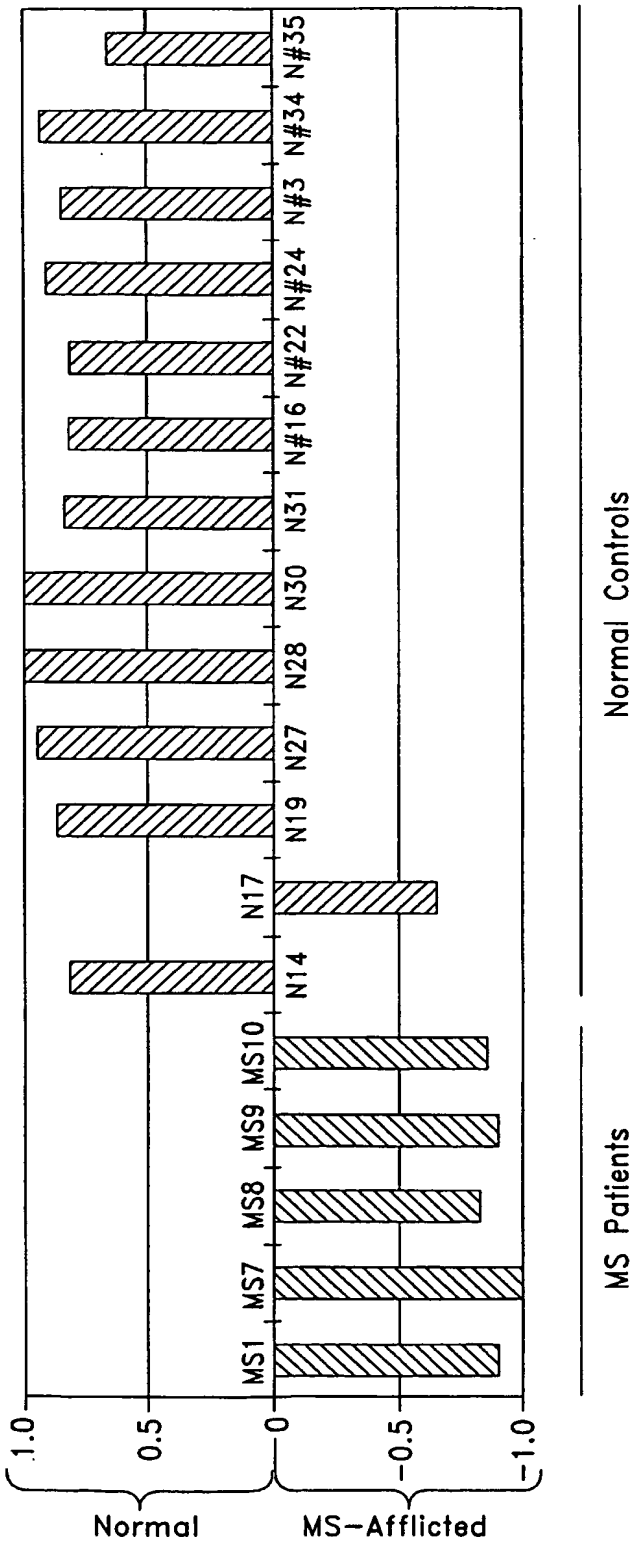
N#35
N#34
N#3
N#24
N#22
N#16
N31
N30
N28
N27
N19
N17
N14
MS10
MS9
MS8
MS7
MS1
Normal Controls
MS Patients
1.0
0.5
0
-0.5
-1.0
Normal
MS-Afflicted

FIG. 2

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US02/09305

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) : C07H 21/04; C12Q 1/68
US CL : 435/6; 536/23.1
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 435/6; 536/23.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Please See Continuation Sheet

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	VAN BOXEL-DEZAIRE et al. "Decreased interleukin-10 and increased interleukin-12p40 mRNA are associated with disease activity and characterize different disease stages in multiple sclerosis" Annals of Neurology, June 1999, Vol. 45, No. 6, pages 695-703, see entire document.	1-26
A	SEDLACEK et al. "Evolutionary conservation and genomic organization of XAP-4, an Xq28 located gene coding for a human rab GDP-dissociation inhibitor (GDI)". Mammalian Genome. October 1994, Vol. 5, No. 10, pages 633-639, see entire document.	1-26

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
23 July 2002 (23.07.2002)	22 AUG 2002
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703)305-3230	Authorized officer Jeanine Enewold Goldberg Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US02/09305

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claim Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claim Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claim Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
Please See Continuation Sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-26, in part, namely gene XAP4

Remark on Protest
☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/09305

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING

Group 1-808, claim(s) 1-26, in part, drawn to methods of diagnosing a subject with multiple sclerosis by comparing expression of a marker listed in Tables 7-10 to normal expression levels or using a panel of markers for expression analysis of multiple sclerosis. It is noted that Groups 1-808 correspond to the 808 markers of Tables 7-10. Therefore, the first mentioned invention is the methods of claim 1 to the extent that they apply to expression analysis of XAP4 (Accession U67322) and panels of at least five markers wherein the panel comprises XAP4. **Group 1, the first mentioned invention, is the invention which will be searched in accordance with PCT Article 17(3)(a).** Additional groups may be elected. For example, if Group 2 is elected, and the proper fees are paid, then claim 1 will be examined to the extent that they apply to methods of expression analysis with respect to OA48-18 (Accession Number AF069250) or a panel of five makers comprising said marker recited in claim 1. Upon election of an invention to be searched in addition to group 1, please identify the number of the gene to be searched in the method claims.

Group 809-1616, claim(s) 27-35, in part, drawn to methods of assessing the efficacy of therapy or compounds by detecting expression of a marker listed in Tables 7-10.

Group 1617-2424, claim(s) 36-38, 54-58 in part, drawn to kits comprising nucleic acids of Tables 7-10 and biochips.

Group 2425-3232, claim(s) 36, 39, 48-50 in part, drawn to kits comprising antibodies which binds with a protein corresponding to the markers of Tables 7-10.

Group 3233-4040, claim(s) 40, drawn to methods of modulating expression using antisense oligonucleotides .

Group 4041-4848, claim(s) 41-42, drawn to methods of modulating expression using a protein of Tables 7-10.

Group 4849-5656, claim(s) 43-44, drawn to methods of modulating expression using an antibody of Tables 7-10.

Group 5657-6464, claim(s) 45, drawn to methods of localizing a therapeutic moiety.

Group 6465-7272, claim(s) 46-47, drawn to methods of screening for test compounds by monitoring expression of protein of Tables 7-10.

Group 7273-8080, claim(s) 51-53, drawn to peptides encoded from markers listed in Tables 7-10.

The inventions listed as Groups 1-89 do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons.

The claims are drawn to detecting expression of numerous genes, in the alternative, as diagnostic of multiple sclerosis. It is well known in the art that various genes are differentially expressed in multiple sclerosis cells such that one may detect a predisposition of MS based upon overexpression of a gene compared to the normal expression level of the gene. The disclosure specifically states that "A study by Whitney and colleagues utilized genetic expression data from one patient to compare MS tissue to non-MS tissue, revealing that about sixty-two genes were differentially expressed, including the Duffy chemokine receptor, interferon regulatory factor-2 and tumor necrosis factor alpha receptor 2." (page 3, lines 5-15). Moreover, it is noted that the art teaches the association of over expression of many genes in MS (Baranzini, J. Immunology Vol 165, pages 6576-6582, 2000). Therefore, there is no special technical feature which links the claims as defined by PCT Rule 13.2, and therefore lack of unity is present between the groups.

Additionally, methods which are comparing expression of one gene are not the same as method comparing expression of a different gene. The expression level of one gene as indicative of MS does not bear any indication of how a different gene will be expressed in cancer cells. The molecules, namely genes, which are relied upon in each of the methods do not share a common structure.

With respect to the markers, namely the polynucleotides and polypepetides, each of the genes provided in Table 7-10 are identified by Accession Number indicating that the sequence of the genes were known in the art. For example, XAP4, Genbank Accession Number U67322, available in 1999 contains both a nucleic acid and a polypepetide sequence. Therefore, the genes themselves are not contributions over the art as exemplified in the Tables themselves.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/09305

The groups comprising polynucleotides, kits, polypeptides, antibodies are additionally drawn to multiple, distinct products lacking the same or corresponding special technical features. The nucleic acids are composed of nucleotides and function in, e.g., methods of nucleic acid hybridization or amplification. These groups are directed to different combinations of nucleic acids which are different from one another and may be employed in different methods. The polypeptides differ in both structure and function from either the nucleic acids or the transgenic organisms. The polypeptides are composed of amino acids linked by peptide bonds and arranged in a complex combination of alpha helices, beta pleated sheets, hydrophobic and hydrophilic domains. The polypeptides also differ in function, e.g., fusion proteins with an enzymatic functions. The antibodies are composed of amino acids linked by peptide bonds, antibodies are glycosylated and their tertiary structure is unique, where four subunits (2 light chains and 2 heavy chains) associated via disulfide bonds into a Y-shaped symmetric dimer. The antibodies function in immunoassays. As products of different sets of Groups differ from each other in structure, function, and effect, they do not belong to a recognized class of chemical compound, or have both a "common property or activity" and a common structure as would be required to show that the inventions are "of a similar nature".
The methods are similarly not the same because they rely upon different products and have different objectives, reagents and results. Therefore, the methods lack unity.

Continuation of B. FIELDS SEARCHED Item 3:
file medline biosis caplus embase scisearch
xap4, xap 4, xap-4 or u67322, hbv associated factor